

DEVELOPING A SYSTEM RESILIENCE APPROACH TO THE IMPROVEMENT OF PATIENT SAFETY IN NHS HOSPITALS

Submitted by

**MICHAEL DERMOT ANDREW
WILLIAMS**

to the University of Exeter as a thesis for the degree of

**DOCTOR OF PHILOSOPHY
IN MANAGEMENT STUDIES**

April 2011

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

..... (Signature)

Abstract

The objective of this thesis is to explore how a systems approach can be used to provide an insight into patient safety in NHS hospitals in England. Healthcare delivers considerable benefits yet there remains a relatively high rate of harm and death for patients through adverse events occurring during the process of treatment. The extant patient safety literature acknowledges the influence of organisational or system factors on patient safety. However, the literature is weak in explaining how system factors affect patient safety. To provide an insight into the interactions within healthcare systems, this research explores the characteristics of NHS hospitals, regarded as complex socio-technical systems, using concepts from resilience, systems, accident and social theory.

A theoretical Safe Working Envelope (SWE) model (Rasmussen, 1997) is developed and contextualised for use in the NHS. The case study field work was carried out in two NHS hospitals during consecutive winter months at times of high demand for inpatient services. A third case study uses secondary data about patient safety failures in the Mid Staffordshire NHS Foundation Trust.

The original SWE model has three failure boundaries. The model is developed by introducing an additional boundary to take account of Government targets. Social theory and system dynamics are used to include the dialectic feedback of social actors and the dynamics of workload. The model depicts the competing pressures, constraints and the workload associated with the need to meet the financial, target, staff workload and patient safety requirements. Three interacting construct sets are explored. These are the constraints within which the system operates, the pressures from the context, and the system dynamics of demand, capacity and decision making. Insights into system behaviours of the hospitals are derived from examining the construct set interactions. The proposition is made that there are five system behaviour archetypes which create the conditions that influence patient safety. The archetypes are derived from the system dynamics and in particular the relationship between reinforcing and balancing feedback loops. The five archetypes are safe practice, drift, tip, collapse and transition towards failure. As hospitals become overcrowded the complexity increases and the reinforcing feedback loops dominate the system and potentially increase the risk to patients. An element of risk arises from staff normalising to the drift in standards of care.

Acknowledgements

Completing this thesis has been a journey shared and made possible by others. My supervisors, Professors Andi Smart and Roger Maull played a huge part in my learning what a PhD needs to achieve; thanks for to them for their patience, understanding and advice. Getting dates in the diaries of supervisors is never easy so special thanks to Francine Carter for her help and assistance.

A special mention and acknowledgement is due to Prof Kieran Sweeney who sadly died in late 2009. Kieran was an informal supervisor who greatly encouraged me in gathering 'rich data' from the stories and encounters with staff during my research. He is greatly missed. I can only hope to have done justice to his enthusiasm for qualitative methods. Thanks also to Prof Charles Vincent for his encouragement.

Thanks are due to those who funded my research. In particular to Sir Ian Caruthers, the Chief Executive and Dr Mike Durkin, the Medical Director of the South West Strategic Health Authority for their encouragement and support to undertake this work. Thanks to Taunton & Somerset NHS Trust who helped me make the transition, due to ill health, from being a Chief Executive to a student.

To many former colleagues in the NHS who have kept me sane by involving me in patient safety projects, with particular thanks to Pete Cavanagh, Peter Aitken and Julie Branter. Thanks especially to the Chief Executives and staff of the case study hospitals for allowing me access.

Thanks to my colleagues in the research group for allowing me to test my ideas with them. For a major part of the time my son Jason was part of the research group. I'm grateful that he did not pose too many difficult questions for his old man! To Prof Peta Foxall at the Peninsula College of Medicine and Dentistry for asking me to work with and teach doctors about leadership and patient safety, which has kept my feet rooted in reality. She also helped by reminding me that the PhD would be finished one day and life would return to normal.

Parts of this thesis have been read by a number of people and particular thanks go to Nigel Walsh and Harry Maddern for helping make it read better than it might have.

Gill my wife has been a solid support as always, ably assisted by Jason, Hannah and Fran plus Millie the dog. There is now hope that my study floor will once more see the light of day.

Table of Contents

Title Page	1
Abstract	2
Acknowledgements	3
Table of Contents	4
Publications	9
List of Figures	9
List of Tables.....	12
Abbreviations	15
Glossary.....	16
Chapter 1 - Introduction.....	18
1.1 Introduction.....	18
1.2 Context of the research.....	18
1.3 Scope of the research	20
1.4 Research aims and objectives.....	23
1.5 Significance of the study.....	25
1.6 Organisation of the thesis.....	26
1.6.1 Chapter 2 – Patient Safety and Systems Thinking.....	28
1.6.2 Chapter 3 – Developing a system resilience approach.....	28
1.6.3 Chapter 4 – Research philosophy, design and methodology	28
1.6.4 Chapter 5 – Developing the Conceptual Model.....	29
1.6.5 Chapter 6 – Investigating the Boundaries of the Safe Working Envelope ...	30
1.6.6 Chapter 7 – Investigating the Gradient and Operating Point.....	30
1.6.7 Chapter 8 – Investigating the Structure and Feedback.....	31
1.6.8 Chapter 9 - Discussion	31
1.6.9 Chapter 10 - Conclusions, Implications and Limitations.....	31
Chapter 2 – Patient Safety and Systems Thinking.....	33
2.1 Introduction.....	33
2.2 Patient Safety – the influence of the ‘system’	33
2.3 System models in the patient safety literature	38
2.4 Systems Thinking.....	41
2.4.1 What is a system?.....	41
2.4.2 System Dynamics.....	43
2.4.3 Decision Making	47

2.4.4	Summary of systems thinking contribution	48
2.5	Summary	49
Chapter 3 – Developing a system resilience approach		51
3.1	Introduction	51
3.2	Concepts from Accident Theory	51
3.3	Resilience	57
3.3.1	Resilience Engineering and the Safe Working Envelope	61
3.4	Synthesis of Resilience and Accident Theory.....	72
3.5	Summary	73
3.6	The research objective.....	74
Chapter 4 - Philosophical Assumptions and Methodology		76
4.1	Introduction	76
4.2	Philosophical Reflexivity	76
4.2.1	The nature of social science and management research.....	77
4.2.2	The Pragmatic-Critical Realist approach	81
4.2.2.1	Appropriateness of the pragmatic-critical realism	83
4.2.3	Structure / Agent debate.....	85
4.2.4	Summary of philosophical assumptions	86
4.3	Philosophical assumptions in the current literature	87
4.4	The research design and methodology	88
4.4.1	Research design.....	89
4.4.2	Justification for the use of case study design.....	91
4.4.2.1	Judging the quality of case studies.....	91
4.4.2.2	The Unit of analysis	94
4.4.3.1	Selection and justification of the case studies.....	94
4.5	Research methods.....	96
4.5.1	Analysis of Documents	99
4.5.2	Semi-structured interviews.....	100
4.5.3	Non-participant observation.....	102
4.5.4	Descriptive statistics.....	102
4.5.5	Questionnaire	103
4.5.6	Causal Loop Diagrams (CLD)	103
4.5.7	Stock Flow Diagrams (SFD).....	104
4.6	Ethical issues.....	104
4.7	Reflexivity.....	106
4.8	Summary	107

Chapter 5 - Contextualising the Conceptual Model.....	109
5.1 Introduction	109
5.1 Developing the conceptual model.....	110
5.1.1 Synthesising the literature	110
5.2 Contextualising the Safe Working Envelope model	113
5.2.1 Policy Context.....	114
5.2.2 Extant literature on targets and performance	116
5.2.3 Developing the model	117
5.3 Applying the model.....	125
5.4 Summary	128
 Chapter 6 – Investigating the Boundaries of a Safe Working Envelope ..	129
6.1 Introduction	129
6.2 Staff views on the priorities of the boundaries.....	133
6.3 Patient Safety Boundary.....	135
6.4 Unacceptable Workload Boundary	144
6.5 Financial Failure Boundary.....	150
6.6 Target Failure Boundary	155
6.7 Summary	162
 Chapter 7 – Investigating the Gradients and Operating Point	166
7.1 Introduction	166
7.2 Event one – Norovirus	168
7.2.1 Response	170
The next section provides evidence of the actions taken in response to the pressures exerted by the different gradient pushing the OP away from breaching the failure boundaries.	171
7.2.1.1 <i>Opening additional inpatient beds</i>	171
7.2.1.2 <i>Increasing the rate of medical outliers</i>	172
7.2.1.3 <i>Implications for the patients</i>	173
7.2.1.4 <i>Keeping ED flowing</i>	176
7.2.1.5 <i>Not cancelling elective patients</i>	177
7.2.1.6 <i>Staff workload</i>	179
7.2.2 The Operating Point and Gradient dimensions	181
7.3 Event two – post Christmas surge in emergency admissions	184
7.3.1 Response- staff workload and patient safety.....	187
7.3.2 Response – decision making hierarchy	188
7.3.3 Consequence – shifting the safety failure marginal zone boundary	189
7.3.4 Analysis of event two.....	191

7.4	Event three – the flow of emergency patients on one day	195
7.4.1	The flow of patients	196
7.4.2	Implications for patients.....	198
7.5	Summary	202
7.5.1	Operating Point	202
7.5.2	Gradients	203
 Chapter 8 – Investigating the Structure and Feedback		 206
8.1	Introduction	206
8.2	The planned design of the structure	207
8.3	Factors that affect the planned stock flow design	210
8.4	Tipping from loose to tight coupling	217
8.5	Increasing the rate of flow out of the medical wards	221
8.6	Conceptual analysis.....	231
8.6.1	‘Structure’	231
8.6.2	‘Feedback’	231
8.6.3	Implications on the OP	233
8.7	Summary	234
 Chapter 9 – Discussion		 235
9.1	Introduction	235
9.2	The development of the Safe Working Envelope model	235
9.2.1	Key points from the literature	236
9.2.2	Insights about the constructs and dimensions of the SWE v3 model	238
9.2.3	Resilience – the buffer of adaptive capacity	243
9.3	Characteristics of the case study hospitals	245
9.3.1	Insights into system archetypes.....	245
9.3	System archetypes that may influence patient safety.....	247
9.3.1	One – Safe Practice	247
9.3.2	Two – Drift towards Failure.....	247
9.3.3	Three – Tip towards Failure.....	249
9.3.4	Four – Collapse towards Failure	251
9.3.5	Five – Transition towards Failure	252
9.4	Proposals for improvement	253
9.5	Summary of the contribution to knowledge.....	255
9.6	Summary	257
 Chapter 10 - Conclusions, Implications and Limitations		 258
10.1	Introduction.....	258
10.2	Summary of Conclusions	258

10.2.1	Conclusions about safe working envelope model.....	258
10.2.2	Conclusions about the systemic characteristics influence on patient safety 260	
10.3	Implications.....	262
10.3.1	Implications for theory.....	262
10.3.2	Implications for policy.....	264
10.3.3	Implications for practice.....	265
10.3.4	Contribution to knowledge.....	267
10.4	Limitations of the research.....	268
10.5	Further research.....	270
10.5.1	System Behaviours.....	270
10.5.2	Operational and patient safety management.....	271
10.5.3	Development of the model.....	271
	References.....	273

Appendices

Appendix 2.1:	Conventions of System Dynamic Models	289
Appendix 4.1:	Documents and Statistical data collected	292
Appendix 4.2:	Research Protocol	297
Appendix 4.3:	List of Interviewees	299
Appendix 4.4:	Interviewee Questionnaire	301
Appendix 4.5:	Information about the research	302
Appendix 4.6:	Consent Form	304
Appendix 4.7:	Ethical Approval	305
Appendix 5.1:	Summary of key points from NHS Operating Framework Documents (2006/07 – 2009/10)	310
Appendix 5.2:	Results of Content Analysis	318
Appendix 6.1:	Extracts from Filed Notebooks	320
Appendix 6.2:	Extracts from coded data	324
Appendix 6.3:	Questionnaire results	360
Appendix 7.1:	Examples of Daily bed information	362
Appendix 8.1	Admission and discharge graphs	365

Publications

Publication 1:	Patient safety: a casualty of target success?	370
Publication 2:	System resilience and patient safety during a bed crisis in an NHS hospital in England	387
Publication 3:	Rebellion against the ‘normal’ to improve safety for patients	398
Publication 4:	Safety, systems, complexity, and resilience: What makes organizations safe?	408

List of Figures

Figure 1.1:	Outline approach of the thesis	22
Figure 1.2:	Overview of research process and chapter content	27
Figure 3.1:	Reason’s Swiss Cheese model (Reason, 1997)	52
Figure 3.2:	Blunt and sharp end of an organisation	53
Figure 3.3:	Two views of ‘resilience’	58

Figure 3.4:	Safe working envelope model	63
Figure 3.5:	Gradients of pressure on the operating point	66
Figure 3.6:	Counter gradient to hold the operating point in safe position	66
Figure 4.1:	Assumptions about the nature of social science	77
Figure 4.2:	Two axis framework depicting schools of management research	80
Figure 4.3:	High level overview of empirical research process	89
Figure 5.1:	Safe Working Envelope (v2) model developed from synthesis of the literature	113
Figure 5.2:	The four boundary SWE for NHS hospitals	118
Figure 5.3:	Four boundary SWE with gradients	119
Figure 5.4	SWE set within the wider context of stakeholder influences	121
Figure 5.5:	Four boundary SWE with SD diagrams to depict system 'structure' within the envelope	122
Figure 5.6:	Basic stock flow diagram of the emergency medical patient pathway into and through a NHS hospital	123
Figure 5.7:	Basic stock flow diagram of the emergency medical patient pathway into and through a NHS hospital with diversion flow from MAU to ED	124
Figure 5.8:	Causal Loop Diagram showing the relationship of the emergency medical patients rate of arrival to a hospital system	125
Figure 5.9	Combination and interaction of construct sets depicted by the SWE model	126
Figure 6.1	Combination and interaction of construct sets depicted by the SWE model	130
Figure 6.2:	Coding hierarchy for 'Safe Working Envelope'	130
Figure 6.3:	Coding hierarchy for 'Actual Design' theme	131
Figure 6.4:	Coding hierarchy for 'Rich Pictures'	131
Figure 6.5:	Relationship of themes to SWE (v3) model	132
Figure 6.6:	Movement of the marginal zone boundary for control of MRSA and C.Diff	137
Figure 6.7:	Movement of the marginal zone boundary for inpatient patient falls	139
Figure 6.8:	Location of the marginal zone boundary for nurses and	146

doctors

Figure 6.9:	The SWE with boundaries dimensions	164
Figure 7.1	Combination and interaction of construct sets depicted by the SWE model	166
Figure 7.2:	The OP breaching the patient safety failure boundary due to the rapid spread of the Norovirus in CS 1	169
Figure 7.3	Movement of the OP following the outbreak of the Norovirus	171
Figure 7.4:	Statistical Process Control Chart - Medical outliers before, during and after sickness virus	175
Figure 7.5:	Elective Admissions 2008 Before, During and After Sickness Virus and 2007 over same period	178
Figure 7.6:	Event one – SWE (v3) with operating point and actions and consequences associated with gradients from each boundary	183
Figure 7.7:	CS 1 Predicted emergency medical admission with control limits with actual admissions Dec 08 – Jan 09	186
Figure 7.8:	Number of medical outliers by day May 08 – Feb 09	186
Figure 7.9:	CS 1 Shifting the marginal zone boundary outwards	192
Figure 7.10:	Event two – SWE (v3) with operating point and actions and consequences associated with gradients from each boundary	194
Figure 8.1	Combination and interaction of construct sets depicted by the SWE model	206
Figure 8.2:	CS 1 Elective Inpatient Admissions, Plan vs Actual, Apr 08 – Sept 09	208
Figure 8.3:	CS 1 Emergency (Non-Elective) Inpatient Admissions, Plan vs Actual, April 08 – Sept 09	209
Figure 8.4:	The planned stock and flow of surgical and medical patients in CS 1	210
Figure 8.5:	CS 1 Daily medical admissions and discharges Dec 08 – Jan 09	211
Figure 8.6:	CS 1 Monday Median Elective and Emergency Admissions vs Discharges by Hour of the Day 2006-07	212
Figure 8.7:	CS 1 Friday Median Elective and Emergency Admissions vs Discharges by Hour of the Day 2006-07	212
Figure 8.8:	CS 1 Variation between Elective and Emergency Admissions vs Discharges April 06 – March 07	213

Figure 8.9:	CS 1 31 Bed Acute Medical Ward: ‘Enhanced’ Occupancy (number of patients on ward per day) in October 2008	215
Figure 8.10:	CS 1 31 Bed Acute Medical Ward: Midnight and ‘Enhanced’ Percentage by day - October 2008	215
Figure 8.11:	Changes (depicted in red) to the planned design of the stock and flow of surgical and medical patients in CS 1 and 2 when medical wards are full	218
Figure 8.12:	CLD of reinforcing loops created in situation where medical ward occupancy is too high to absorb variation in demand. (Stocks in boxes)	219
Figure 8.13:	CLD with balancing loops of increasing rate of flow out of medical wards	222
Figure 8.14:	CS 1 Number of Medical Outliers Dec – June 2006-2009	224
Figure 8.15:	CS 2 Number of Outliers Jan 09 – Jan 10 (with missing data points)	224
Figure 8.16:	CS 1 Length of Stay of Medical Patients by Discharge Ward	225
Figure 8.17:	Reinforcing feedback loops generated by ‘medical outliers’	227
Figure 8.18:	Feedback loops generated by transfer of medical patients to surgical wards	230
Figure 9.1:	The constructs of the SWE v3 model	242
Figure 9.2	Overview of the system characteristics and behaviour archetypes	246

List of Tables

Table 1.1:	The contrast between human error and system vulnerability to failure	21
Table 2.1	Overview of main themes in the patient safety literature	35
Table 2.2	Variety of ‘system’ concepts found in the literature	37
Table 3.1:	Dimensions of the SWE constructs	71
Table 4.1:	Categorising the implications for research of the four paradigms	79
Table 4.2:	Case study tactics to meet the four design tests	92

Table 4.3:	Summary of sources of data and data collection methods	97
Table 4.4:	Summary of data sources and their application to themes	98
Table 4.5:	Summary of data sources used to establish the characteristics of the SWE constructs	98
Table 4.6:	Case study 1 number of interviewees by profession and embedded level in the hospital	101
Table 4.7:	Case study 2 number of interviewees by profession and embedded level in the hospital	101
Table 5.1:	Mean scores by boundary type from content analysis of Department of Health ‘Operating Frameworks’ for the NHS in England 2006/07 – 2009/2010	115
Table 6.1:	Dimensions of the SWE ‘boundaries’ constructs derived from literature	132
Table 6.2:	CS 1 - Ranking of the highest priority for different organisational levels	133
Table 6.3:	CS 2 - Ranking of the highest priority for different organisational levels	133
Table 6.4:	Dimensions of the SWE ‘Patient safety failure boundary’ construct with examples from the data	143
Table 6.5:	Dimensions of the SWE ‘Unacceptable workload boundary’ construct with examples from the data	150
Table 6.6:	Dimensions of the SWE ‘Financial failure boundary’ construct with examples from the data	155
Table 6.7:	Dimensions of the SWE ‘Financial failure boundary’ construct with examples from the data	162
Table 7.1:	Dimensions of the SWE ‘operating point’ and ‘gradients’ constructs derived from literature	167
Table 7.2:	Summary of the dimensions of the ‘operating point’ during Event 1	184
Table 7.3:	Summary of the dimensions of the ‘gradients’ during Event 1	184
Table 7.4:	Medical Emergency Admissions and Medical Outliers (29 th Dec 2008 – 14 th Jan, 2009)	185
Table 7.5:	Summary of the dimensions of the ‘operating point’ during	195

the surge in demand event

Table 7.6:	Summary of the dimensions of the ‘gradients’ during the surge in demand event	195
Table 7.7:	Summary of the dimensions of the ‘operating point’ during Event 3	201
Table 7.8:	Summary of the dimensions of the ‘gradients’ during Event 3	201
Table 8.1:	Explanation of loops 5 – 9 in Figure 8.16	228-9
Table 8.2:	Dimensions of the ‘structure’ and ‘feedback’ constructs of the SWE (v3)	232
Table 9.1:	Concepts from accident theory and their relationship to the SWE v3 model constructs.	238
Table 9.2:	Summary of the contribution to knowledge from the development of the SWE model	240-1
Table 9.3	Measurable indicators of the constructs to provide an insight into state of resilience of a hospital	244

Abbreviations

A & E	Accident and Emergency Department (also known as ED)
BTCU	Borderline tolerated conditions of use
C.Diff	Clostridium difficile (infection)
CEO	Chief Executive Officer
CLD	Causal Loop Diagram
DoO	Director of Operations (also known as Chief Operating Officer - COO)
CQC	Care Quality Commission
CS	Case Study
ED	Emergency Department
EMU	Emergency Medical Unit
FT	Foundation Trust
GP	General Practitioner
HC	Healthcare Commission
HRT	High Reliability Theory
HRO	High Reliability Organisation
ITU	Intensive Care Unit
MAU	Medical Assessment Unit
MRSA	Meticillin Resistant Stapylococcus Aureus (infection)
NAT	Normal Accident Theory
NHS	National Health Service
OP	Operating Point
SAU	Surgical Assessment Unit
SFD	Stock Flow Diagram
SD	System Dynamics
SWE	Safe Working Envelope
UK	United Kingdom
USA	United States of America
WHO	World Health Organisation

Glossary

A&E	Accident & Emergency department of hospitals where emergency patients are first seen and treated before admission or discharge. Also known as the Emergency Department (ED).
Accident Theory	Theories and concepts derived from the study of accidents and disasters
Adverse event	“an unintended injury or complication resulting in prolonged hospital stay, disability at the time of discharge or death caused by healthcare management rather than by the patient’s underlying disease process.” De Vries E.N. et al (2008).
Boundaries	Boundaries of the Safe Working Envelope (SWE) that depict the constraints within which the hospital works.
Buffer	The capacity to absorb and or adapt to disruption or continuous stress. The SWE model depicts the buffer capacity as the ‘marginal zone’.
Complex socio-technical system	Systems with a large number of elements (technical and social) which interact in a non-linear way with feedback loops and are open to their environment.
Department of Health	Government department responsible for developing health policy and the management of the NHS in England.
Feedback Loop	Feedback from one element of a system to another – often non linear.
Gradient	Within the SWE model the gradient depicts the pressure being exerted on the Operating Point of the system.
Marginal zone	Within the SWE model the area close to the failure boundary depicted by a marginal zone boundary (dotted line).
National Health Service	Publically funded and run healthcare organisation in the UK.
Operating Point (OP)	Within the SWE model the OP depicts the performance of the system in relation to boundary constraints.
Patient safety	There are few definitions found in the literature. The one used is “The avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of healthcare.” (Vincent, 2010, p.31)
Redundancy	Sources which are not used except to cover gaps or failure in normal operations.

Resilience	‘...the ability of an organisation (system) to keep, or recover quickly to, a stable state, allowing it to continue operations during and after a major mishap or in the presence of continuous significant stresses.’ (Wreathrall, 2006)
Social Structure	“...patterns of institutions and relations are the results of actions on the part of individuals who are endowed with the capacities or competencies that enable them to produce these structures by acting in organised ways. ...Embodied structures are found in the habits and skills inscribed in human bodies and minds. These embodied structures allow them to produce, reproduce and transform their institutional and relational structures.” (Scott, 2001. p.84)
Structure (system dynamics)	That structure consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it.’ (Sterman, 2000, p.107)
System	Checkland (1981) suggests that a system: ‘...embodies the idea of a set of elements connected together which form a whole, this showing the properties which are properties of the whole, rather than properties of its component parts’ (p.3).

Chapter 1 - Introduction

1.1 Introduction

This chapter sets the context, rationale and scope of this research and provides an overview of the contribution to knowledge. There are two main motivations for conducting the research. The first is the previous personal experience of the researcher working as a manager in National Health Service (NHS) hospitals where significant harm and deaths of patients occurred. The second is to take a systems perspective to contribute to the patient safety literature. Undertaking this study has been a journey of both discovery and consolidation. It has been a process of discovery by finding new ideas in the literature and during the research process. The consolidation has occurred through times of reflection on previous experience in the light of new learning from the literature. The thesis reflects the combination of academic rigour and practical requirements, borne out of experience, to improve safety for patients. The structure of the thesis is presented and short summaries provided of the content of each chapter.

1.2 Context of the research

Healthcare is one of the basic needs in any society to help save life, overcome disease and relieve suffering. The provision of healthcare faces a number of challenges. In developed countries these include increasing costs, an expanding older population, new disease patterns associated with wealth, and an increasingly complex and sophisticated means of delivering treatment through teams of professionals using new techniques and equipment. Alongside these challenges are the ethical principles underpinning the delivery of healthcare. The ethical concepts of nonmaleficence (do not cause harm) and beneficence (contribute to a persons welfare) are key underlying principles for clinicians (Beauchamp and Childress, 1989). Patients expect that the delivery of treatment will be conducted safely and that the process of care will not endanger their wellbeing beyond the disease process itself.

Florence Nightingale was one of the first to note that hospitals can be dangerous places for patients. She pointed out that the rate of death in certain hospitals was far higher than should be expected. Introducing her 'Notes on Hospitals' in 1863, she wrote:

“It may seem a strange principle to enunciate as the very first requirement in a Hospital that it should do the sick no harm. It is quite necessary, nevertheless, to lay down such a principle, because the actual mortality in hospitals, especially in those of large crowded cities, is very much higher than any calculation founded on the mortality of the same class of diseases amongst patients treated out of hospital would lead us to expect.” (Sharpe and Faden, 1998)

The term ‘iatrogenic disease’ was introduced in the 1920s as a term to describe the harmful effects of medical treatment (Vincent, 2010). Writing in 1970s, a critic wrote of iatrogenic disease having reached epidemic proportions (Illich, 1997). The book commences with the statement: “The medical establishment has become a major threat to health.” The threats were known but few studies were undertaken into the scale and reasons for the rate of harm associated with medical treatment. One of the first prospective studies conducted was in the early 1960s at Yale University Hospital by Elihu Schimmel (Vincent, 2010). The research, involving a 1000 patients, found that ‘20% of the patients admitted to the medical wards suffered one or more untoward episodes and 10% had a prolonged or unresolved episode.’ (quoted in Vincent, 2010). Sixteen fatalities were noted in the study. In the early 1980s research sought to reassess the situation (Steel et al., 1981). The study involved 815 patients and 36% were found to have suffered an iatrogenic illness with 9% classified as threatening to life or causing major disability.

The lack of wide reaching research in this area during this period is claimed to be negligent (Vincent, 1989). In the UK attitudes changed after the events at Bristol Royal Infirmary where paediatric heart surgery death rates were judged to be excessive. It took a number of years for the problem to come to light and the subsequent inquiry shone a powerful light into, not just the performance of the surgeons but also the culture and wider systemic influences (Kennedy, 2001). In the USA the publication by the US Institute of Medicine’s report ‘To err is human’ awoke professional and public interest in the scale of preventable harm and death occurring in healthcare and hospitals in particular (Kohn et al., 2000). The report claims that more people die from medical errors in the USA in any one year than from car accidents, breast cancer or AIDS.

The scale of harm is not easy to define or measure. Researchers employ different methods with a range of strengths and weaknesses. Studies where the researchers

retrospectively review the patient case notes for adverse events have been used in a number of countries (Vincent, 2010). Whilst researchers set different criteria as to what constitutes an adverse event, the typical rate appears to be between 8 – 12% of patients admitted to hospital suffer one or more adverse events (de Vries et al., 2008). In the UK a government report on patient safety suggests the rate of 10% of inpatients suffer an adverse event in NHS hospitals (Department of Health, 2000b). The report emphasises the need to learn from incidents and proposes setting up a reporting and learning system (Department of Health, 2001). A later review indicates that progress is not as good as expected and further reforms of the structure supporting patient safety in the NHS is needed (Department of Health, 2006c). In the USA, a review ten years after the Institute of Medicine report, highlighting the scale of the problem, found improvement to be ‘frustratingly slow’ (Leape et al., 2009).

In summary, it is clear that whilst healthcare can provide significant benefit there is a considerable rate of preventable harm and death. It is only comparatively recently that governments, professions and academia have given the subject the attention it deserves. There is a growing realisation that improving patient safety is not easy despite seeking to learn about safety through reporting mechanisms and from other industries. The complexities of the process of treatment alongside the pathway of disease is very different to keeping an aeroplane from falling out of the sky (Vincent, 2010, Walshe and Boaden, 2006). The lack of visibility to the wider society of the scale of the problem in healthcare, compared to the obvious causalities of transport crashes for example, means that there is not the same emotional and political drive to improve.

1.3 Scope of the research

Keeping patients safe is a complex and large scale undertaking (Vincent, 2010). This thesis seeks to address the broad conceptual challenge of how to improve the understanding of organisational and operational influences on patient safety. The research brings a systems thinking approach to the task.

The assumption is often made that things go wrong due to ‘human error’. It is easy to blame the individual clinician as being a ‘bad apple’ (Dekker, 2007). However, there is a growing realisation of wider system issues that contribute to failures, which requires a different approach to learning and improvement (Vincent et al., 1998, Cook et al., 1998,

Woods et al., 2007, Woods et al., 2010, Vincent et al., 2004, Vincent, 2004, Dekker, 2011). The move towards a more holistic approach to patient safety is summarised by Woods et al (2010) in the contrasting two approaches to failure.

First stories – individual ‘bad apple’ theory	Second stories – system vulnerability
Human error (by any other name: violation, complacency) is seen as the cause of failure	Human error is seen as the effect of systemic vulnerabilities deeper inside the organization
Saying what people should have done is a satisfying way to describe failure	Saying what people should have done does not explain why it made sense for them to do what they did
Telling people to be more careful will not make the problem go away	Only by constantly seeking out its vulnerabilities can organizations enhance safety

Table 1.1: The contrast between human error and system vulnerability to failure (adapted from Woods et al, 2010)

The understanding is that the failure to keep patients safe is connected to the wider system, which is influenced by its context (Katz and Kahn, 1966, Cummings et al., 2001). Taking the context into account widens the consideration of factors that need to be considered when seeking to make improvements (Boaden and Burnes, 2009). An underlying assumption made in this thesis is that systems are made up of interacting parts from which safety or failure is an emergent process (Checkland, 1981, Hollnagel, 2004, Woods et al., 2010, Boaden and Burnes, 2009, Forrester, 1969, Dekker, 2011). Dekker (2011) argues that the theories used in the extant literature and practice to understand why failures occur, are not up to the task of fully explaining how complex socio-technical systems create unforeseen consequences. Current theories rely on reductionist and linear ideas to help explain what goes wrong. However, ‘the world is not linear’ (Meadows, 2008). To improve our understanding of how to improve patient safety, it is therefore necessary to explore the characteristics of the dynamic interactions of the parts and their contribution to the behaviour of the whole. To take account of the requirement to investigate the interactions, this thesis takes a systems thinking approach. The relationship of the systems approach adopted in this thesis to the extant literature is outlined in Figure 1.1.

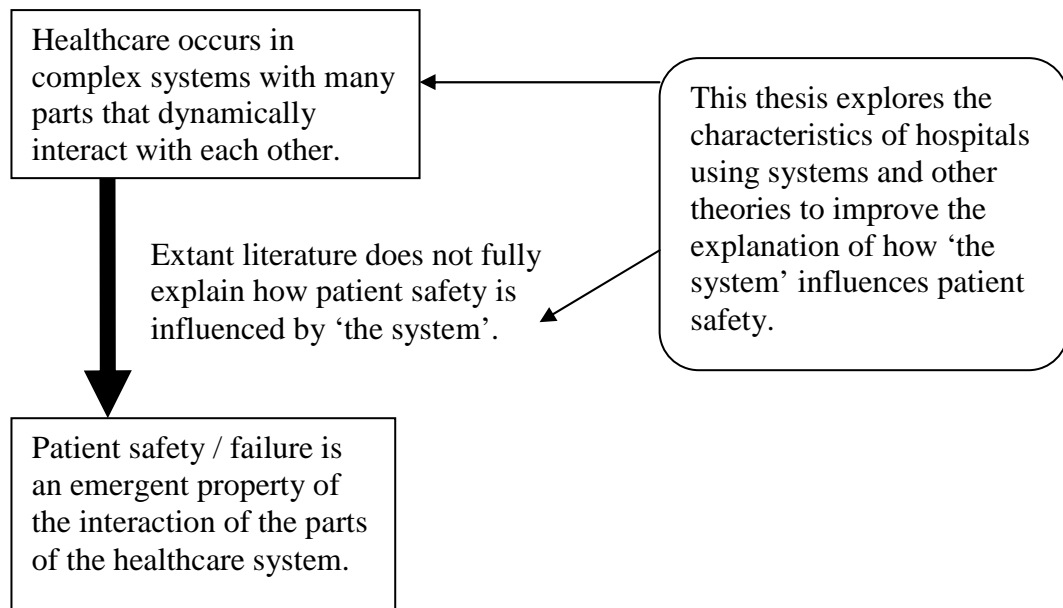


Figure 1.1: Outline approach of the thesis

Taking a systems approach to improve our understanding about patient safety is a broad canvas. There are multiple perspectives on what constitutes patient safety, such as debates about where it sits in the quality agenda and how improvements can be made (Walshe and Boaden, 2006). This thesis focuses on a high level view of the characteristics of hospitals rather than examining specific components.

The definition of patient safety used for this thesis is:

“The avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of healthcare.” (Vincent, 2010, p.31)

The extant patient safety literature widely recognises the systemic influences on patient safety. However, it is argued in Chapter 2 that the system frameworks and models found in the literature have limited explanatory power. In particular the patient safety literature does not engage with systems theory or the nature of complexity to fully explore the influence of the interaction of the parts of a system on safety. Systems theory can assist in developing theoretical models to examine the dynamic interactions that occur between the parts of the system and thereby strengthen the theoretical basis of the literature on safety failures (Dekker, 2011).

Therefore, it is argued that exploratory research is required to examine the characteristics of healthcare systems in relation to patient safety. The resilience literature provides an appropriate theoretical foundation as it links systems theory with

safety. A model, derived from the literature, is developed and used as the basis to conduct the exploratory research.

1.4 Research aims and objectives

The overall aim of the research is to explore some of the characteristics of complex socio-technical healthcare systems that can influence patient safety. The literature review, presented in Chapter 2 and 3, identifies a number of gaps in knowledge to provide a narrower focus. The most significant gaps in the patient safety literature are the limitations of the current theoretical models. First, extant models are weak at being able to explain the influence from the interaction of the parts in complex socio-technical systems, such as hospitals, on the safety of patients. This may partly be due to the apparent lack of systems thinking being applied to the problem. Second, they are limited in using insights from accident theory to enhance the theoretical framework of patient safety. To provide a practical focus within the literature and for the empirical research, hospitals are used as an example of a complex socio-technical healthcare system. The gaps identified provide the basis for the objective of the research, which is:

To explore, in NHS hospitals, how a systems approach can inform the development of patient safety theory.

The limitations of the current theoretical models are addressed through a process of synthesising concepts and models from the systems theory and resilience literature. The research takes an empirical theory development approach (Meredith, 1998). Such an approach leads to the development of propositions for testing in future research.

To operationalise the research objective a theoretical Safe Working Envelope (SWE) model is developed from a synthesis of the patient safety, systems, accident theory, and resilience literature, which is presented in Chapter 3. The foundation of the model is the SWE proposed by Rasmussen (1997). The envelope model takes account of safety within the competing pressures of complex socio-technical systems. A resilient system is one that can operate within the boundaries of the envelope in the face of disruption. It is argued that resilience is part of the same continuum as vulnerability and hence appropriate for exploring patient safety. The model is further developed and contextualised in Chapter 5. A key assumption made in the research, supported by the literature, is that overcrowded hospitals, and those under workload pressure, are

associated with patient safety failures (Fatovich et al., 2005, Sprivulis et al., 2006, Trzeciak and Rivers, 2003, Wears et al., 2008, Weissman, 2007, Cameron, 2006, Kc and Terwiesch, 2009). Therefore, the empirical investigation explores hospitals at a time of pressure and overcrowding, to gain insights into some of the characteristics of the interactions and their impact on system behaviours that may influence patient safety.

The empirical phase of the study focuses on NHS hospitals in England. The structure of the NHS across the UK is different in each nation. In England there are particular policies and structures in place that influence the operating conditions and characteristics of hospitals (Klein, 1995, Ham, 2009). The researcher has considerable experience of working in NHS hospitals in England, providing both a depth of knowledge but also the potential for bias. The potential for bias is recognised and methods are employed to minimise its influence. Hospitals are chosen to study rather than the wider healthcare system as they provide clearly definable organisational systems that interact with a wider context. Most of the extant patient safety literature focuses on hospitals, although it is argued that the theoretical systems perspective is weak.

Three NHS Trust hospitals form the case study part of the research. Case study design, the case selection process and collection methods are justified in Chapter 4. The two case studies (CS 1 and 2) where primary data was collected are chosen to provide contrasting internal characteristics, whilst having the same local Primary Care Trust (PCT) and Strategic Health Authority (SHA). The selection reduced the external differences and allowed the research to focus on how the internal interactions and systemic characteristics responded to disruptive events (perturbations) and periods of continuous stress. The hospitals were studied over the winter period to provide examples of difficult operational situations. The examples amplify the competing pressures, the ability of hospitals to adapt to them, and display the vulnerable and resilient characteristics.

The third case study (CS 3) is an example of significant patient safety failure (Healthcare Commission, 2009a). Mid Staffordshire Hospital NHS Foundation Trust is chosen as there are publicly available inquiry reports into how the competing demands of saving money, meeting targets, staffing the hospital and patient safety were managed.

The third case study is used to supplement the primary data collected from the other two studies.

The case study data and analysis is presented in Chapters 6 – 8 and the conclusions drawn about the characteristics of NHS hospitals in England and how they influence patient safety are discussed in Chapter 9. An overview of the research process and chapter contents is present in Figure 1.2 in Section 1.5 below.

1.5 Significance of the study

The contribution to knowledge is in three areas. The first is how the system characteristics influence patient safety, second, the proposals for improvement and third, the development of the literature on resilience and the SWE model (Rasmussen, 1997) in particular. These are summarised in Section 9.5

First, is the contribution to the patient safety and resilience literature through the incremental development of the SWE model. The development of the model through the synthesis of the literature provides insights in the exploration of some of the interactive characteristics found in NHS hospitals that influence patient safety. The case study evidence adds to the literature that uses the SWE model. For example, empirical evidence of both the chronic and sudden exhaustion of adaptive capacity to prevent patient safety failure is found in the case studies, provides clearer insight as to how that occurs.

Second, from the analysis of the case studies three interacting construct sets are identified which create key system characteristics. The interactions of the construct sets create emergent system behaviour effects, some of which are problematic for patient safety. Five archetypes of system behaviour are identified, four of which increase the risk to the safety of patients. The construct sets and system behaviour archetypes adds a different perspective to the extant patient safety literature. In particular the analysis using the synthesis of concepts from systems thinking, resilience and accident theory provides a new approach to considering how the ‘system’ influences patient safety.

Third, theory building propositions are made about how to reduce the risk associated with the system archetypes. A key outcome of this exploratory research is that tight

coupling between the parts of hospitals and the lack of buffer capacity makes them vulnerable to patient safety failure. Hence the proposition is made that hospitals must be able to accommodate the net difference in rates of flow into and out of the hospital to avoid medical patients being accommodated on surgical wards. The case studies show evidence of the gradual deterioration of standards of care that can occur when hospitals are under continuous pressure to meet targets and save money. The proposition is made that the safety standards have to be made more explicit and judged against high performing organisations and not internal past performance. Building on the extant literature, which suggests that aspects of resilience can be measured, a range of variables are proposed (set out in Table 9.3), to measure the adaptive capacity of hospitals relating to finance, targets, staff workload and patient safety. It is argued that the measures provide a means to assess the resilience of a hospital and the balance made by decision makers between production pressures and the need to protect patients and staff.

1.6 Organisation of the thesis

This thesis is presented in ten chapters. The layout of the thesis is illustrated in Figure 1.2. A short summary of the content of each chapter is set out below.

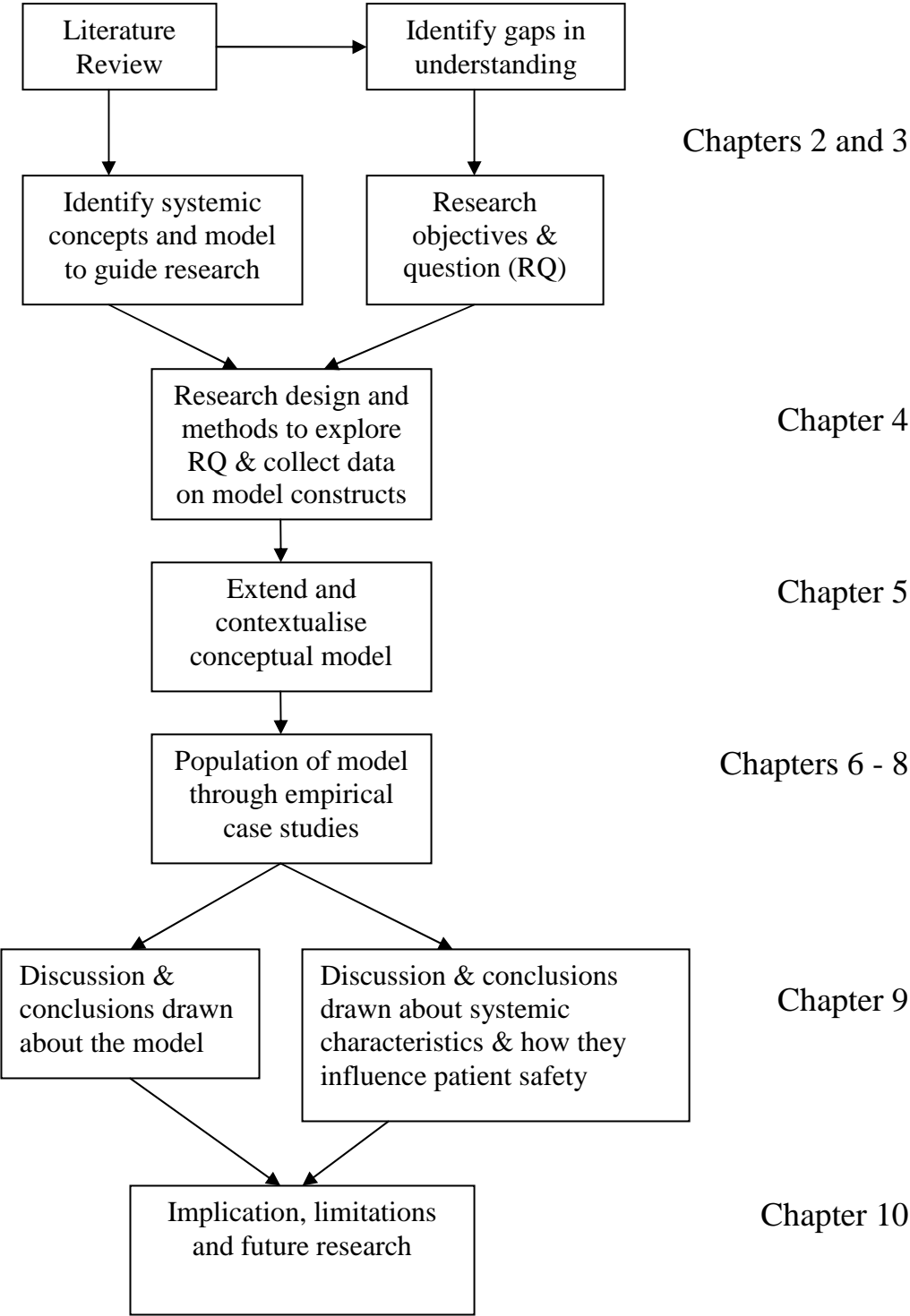


Figure 1.2: Overview of research process

1.6.1 Chapter 2 – Patient Safety and Systems Thinking

This chapter explores the current literature relating to patient safety. Concepts in patient safety are informed from a number of disciplines and perspectives. It is argued that healthcare operates within a complex socio-technical system. The influence of ‘organisational’ and or ‘system’ factors on patient safety is widely acknowledged. Models and frameworks that take account of system factors are found in the literature. However, there is little engagement with systems theory to assist in explaining how the dynamic and interactive characteristics of healthcare systems influence patient safety. It is concluded that the extant literature does not have a clear understanding of the key characteristics of a ‘system’. Therefore, the second part of the chapter reviews the systems thinking literature. The review defines what a ‘system’ is and draws out from the literature certain features of complex socio-technical systems. Complexity theory provides some insights. However, it is argued that system dynamics (SD) concepts incorporate key constructs of healthcare systems. These include the flow of work, the linkage (coupling) and feedback between the parts, the role of decision makers and the resulting behaviour of the whole.

1.6.2 Chapter 3 – Developing a system resilience approach

Having recognised the lack of systems theory in the extant patient safety literature, this chapter reviews the wider accident theory literature. The underlying tension between production pressures and safety are noted along with the dominance of barrier and component failure approach to safety failure. Insights from High Reliability Theory and Normal Accident Theory are derived before considering the concept of resilience.

System-wide and organisational perspectives of safety are explored in the resilience engineering literature. A system level SWE model (Rasmussen, 1997) is identified and developed to explore the influence of system characteristics on patient safety. The review of core themes and the subsequent synthesis of the literature inform the development of the SWE model to operationalise the research objective. The gaps in knowledge are summarised and the research objective is identified.

1.6.3 Chapter 4 – Research philosophy, design and methodology

This chapter sets out some of the philosophical assumptions of the thesis, the research design and the methods employed to investigate the research objective. The chapter has six sections. The basis for understanding what constitutes knowledge and reality is

explored in Section 4.2. The argument is made that taking an objective stance is not possible in complex socio-technical situations. Equally to take a subjective position leaves us with an entirely relative position. The dualistic subjective or objective approach to knowledge and reality is limited and therefore, the dialectic position of 'pragmatic critical realism' (Johnson and Duberley, 2000) is followed.

Section 4.3 looks briefly at the philosophical assumptions found in the literature examined in Chapter 2 and 3. Section 4.4 addresses the methodological options available. It is argued that case studies provide the best design to explore the concepts identified in the SWE model. The strengths and weaknesses of using case studies are set out. The selection criteria and details of the cases chosen are given. Section 4.5 sets out the data collection and analysis methods used and how they are employed. The ethical issues are identified the position and influence of the researcher in the research process is acknowledged.

1.6.4 Chapter 5 – Developing the Conceptual Model

This chapter presents the development and contextualisation of the SWE model. The weaknesses of the model identified in Chapter 3 are addressed by synthesising concepts from SD and social theory. The model is then contextualised for use in research within the NHS.

The contextualisation is achieved through a hermeneutical and content analysis of NHS policy documents and the examination of the literature about performance measurement and targets. The results show the importance of performance targets for NHS hospitals, which necessitates the inclusion of an additional 'target failure' boundary construct within the model. In addition to the 'boundary' construct, four other constructs are identified. The additional insights into the system constructs derived from the SD literature are illustrated. The chapter concludes by arguing that the SWE constructs and the concepts from the safety literature create a theoretical version three of the SWE model, which can be used to explore the systemic characteristics of NHS hospitals and how they influence patient safety. The SWE (v3) constructs are identified as the envelope 'boundaries', 'gradients', 'operating point', 'structure' and 'feedback'.

1.6.5 Chapter 6 – Investigating the Boundaries of the Safe Working Envelope

This chapter explores the ‘boundary’ construct of the SWE (v3) model developed in Chapter 5. The boundary construct depicts the constraints within which the system is designed to operate. Each conceptual failure boundary is considered in terms of how staff themselves describe or explain the competing constraints that they work with. The data from the interviews and observations (CS 1 and 2) and Inquiry Reports (CS 3) is supplemented with analysis of documentation, such as reports and policy papers from the Department of Health. The hospitals are studied during times of high demand for their inpatient services where staff manage complex interactions to keep the hospital functioning. Such situations provide insights into how the prioritisation of competing interactions is managed. While the boundaries are not an observable phenomenon in themselves, observable data relating to the articulation, measurement, and prioritised actions of staff relative to each boundary theme is studied.

1.6.6 Chapter 7 – Investigating the Gradient and Operating Point

This chapter explores the characteristics of the ‘gradient’ and the ‘operating point’ (OP) constructs of the SWE (v3) model. The OP depicts the operating conditions of the system in relation to the boundary constraints. There is a gradient related to each boundary. The gradients depict the pressure exerted on the OP from the internal and wider context to keep it away from the related boundary. The exploration of the characteristics is conducted through the analysis of three events from case study CS 1 and 2. The first event from CS 1 is an outbreak of a sickness virus in the hospital which closed a number of wards for just over a week. Conceptually this event depicts a sudden perturbation when the OP breaches the patient safety boundary, which generated actions and interrelationships between the competing gradients. The second event, from CS 1, is a peak in emergency demand that lasted more than two weeks, which created significant operational problems. This event illustrates ‘continuous stress’ on the hospital, the shift of the marginal zone boundary and the movement of the OP towards the patient safety failure boundary. The third event, from CS 2, is the flow of emergency patients through the Emergency Department (ED) and Medical Assessment Unit (MAU). The event provides an insight into how the dynamics of the gradients and OP movement, at a hospital system level, impact at the micro patient experience.

1.6.7 Chapter 8 – Investigating the Structure and Feedback

This chapter explores the ‘structure’ and ‘feedback’ constructs of the SWE (v3) model. These two constructs are used to gain insights into the characteristics of the dynamics that occur in the SWE, which influence the OP. SD is used to investigate the ‘structure’ and ‘feedback’ loops that occur in the case study hospital systems.

A Stock Flow Diagram (SFD) is used in section 8.2 to illustrate the planned design of the patient flow in and out of CS 1 and 2 hospitals. The planned design is then amended in Section 8.3 to reflect the reality of the situation when the hospitals face either a perturbation, or the continuous stress of high levels of demand for inpatient beds. Two changes in the design are shown. Causal Loop Diagrams (CLDs) are used to show the feedback loops that result from the increased coupling between the parts of the system when the stocks become full and the flows change. From the analysis of the data presented in the diagrams the importance of bottleneck resources to accommodate the net difference in the flow rates of patients in and out of hospital is identified.

1.6.8 Chapter 9 - Discussion

This chapter summarises and discusses the results, assessing them in relation to the literature and draws out the contribution to knowledge. Insights into the characteristics of the case study hospitals are developed using the constructs of the SWE model and concepts from accident theory. The proposal is made that there are three construct sets, which interact to create five archetype system behaviours. The behaviours interact with the resilient capacity of the hospitals to influence the safety for patients.

The development of the current patient safety literature that uses the Ramussen (1997) SWE model is set out. Propositions are made about the key points of intervention to improve the resilience of the case study hospitals and potential measures of adaptive capacity are identified. The contribution to knowledge is summarised.

1.6.9 Chapter 10 - Conclusions, Implications and Limitations

This chapter summarises the conclusions about how the characteristics of the NHS hospitals examined influence patient safety. Conclusions about the development of the SWE (3) are also made. The implications of the conclusions for theory, policy and

practice are presented. The limitations of the research are noted and proposals for future research are suggested.

Chapter 2 – Patient Safety and Systems Thinking

2.1 Introduction

The aim of this chapter is to explore the current literature relating to patient safety and systems thinking. A feature of patient safety literature is that, unlike tightly defined academic disciplines with an accepted core literature, concepts are informed from a number of disciplines and perspectives. Section 2.2 identifies the concepts and perspectives and determines pertinent themes for further exploration. Specifically, system-wide and organisational perspectives of patient safety are explored in Section 2.3. The literature emphasises the influence of the ‘system’ on patient safety. However, the characteristics of systems that impact on safety, such as the dynamic interaction of the parts, are not well explored. It is therefore argued that the system frameworks found in the extant patient safety literature lack explanatory power. Patients are treated in often large and complex organisations which display features of complex socio-technical systems. Those features need to be included in any explanation of how the ‘system’ influences patient safety.

To develop the explanation a review of the systems thinking literature is presented in Section 2.4. This defines what a ‘system’ is and draws out from the literature certain features of complex socio-technical systems. Complexity theory provides some insights. However, it is argued that SD concepts incorporate key characteristics of healthcare systems. These include the flow of work, the linkage (coupling) and feedback between the parts, the role of decision makers and the resulting behaviour of the whole. Connections between systems thinking and the wider safety literature are then explored in Chapter 3.

2.2 Patient Safety – the influence of the ‘system’

The publication of seminal reports highlighted the scale of the unintentional harm and death occurring within healthcare organisations in the USA (Kohn et al., 2000, Brennan et al., 1991, Leape, 1994). The findings presented in these reports initiated a number of similar studies in other countries, confirming that it is an international problem (Wilson et al., 1995, Baker et al., 2004, Davis et al., 2001). Within the UK, the Department of

Health issued reports drawing on the international evidence in an effort to galvanise improvement in patient safety within the NHS, (Department of Health, 2000b, Department of Health, 2001).

Patient safety is not only wide ranging in terms of international relevance, but it is also broad conceptually. For example, in a paper developing 'A Taxonomy of Error in Health Policy' Joyce et al (2005), refer to the phenomenon of 'medical error', but also mention a wider range of phenomena such as 'adverse events' 'safety in the healthcare system' and 'patient safety'. It is possible that 'patient safety' is an umbrella term under which sits a number of approaches. The Report 'To Err Is Human: Building a Safer Health System' (Kohn et al, 2000) sets out definitions of 'accident' and 'error' and then states that 'safety is defined as freedom from accidental injury.' (p.58) It continues by making the point that: 'Ensuring patient safety, therefore, involves the establishment of operational systems and processes that increase the reliability of patient care.' This quote raises the question about the meaning of the word 'reliability' in this context, which opens a debate beyond the scope of this research. The Report also acknowledges that safety is only one aspect of quality in healthcare.

As noted in Chapter 1, Vincent (2010) defines patient safety as being: "The avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of healthcare." This definition is used in this research as it relates directly to the 'process of healthcare' and includes the broader aspects of prevention, avoidance and amelioration rather than just 'freedom from accidental injury' (Kohn et al , 2000). However, there is often a debate in the literature as to what constitutes adverse events (Brennan et al., 1991, Chang et al., 2005, de Vries et al., 2008, Leape, 2002, Vincent et al., 2001, West, 2000). A question can therefore be asked where is the line drawn between an avoidable adverse outcome and the outcome generated by the risks associated with treatment? It is therefore acknowledged that the definition used from Vincent (2010) may change in the future as the literature develops.

The breadth of factors in patient safety is reflected in the literature and the academic disciplines that inform both theory and practice. Examples of practice range from, but are not limited to, the design of packaging and equipment, the management and prevention of healthcare acquired infection, and the reduction of drug errors, through to understanding organisational safety culture. The major academic disciplines that

contribute to the literature are: medicine; applied psychology; engineering; management and sociology. Examples of the main themes found in the literature are provided in Table 2.1. Each theme has been categorised using a taxonomy of approaches to medical error proposed by Joyce et al (2005). It is suggested that such a taxonomy is ‘emergent’ but can provide an insight into the ‘rationales/ideologies that underlie the various approaches found in the literature.

Theme	Examples of the literature	Categorisation after (Joyce et al., 2005)
Ergonomics: design of equipment and training of staff from a human factors perspective	(Carayon et al., 2007, Carayon et al., 2006, Friesdorf et al., 2007, Tonks, 2008, Lowe, 2006, Flin et al., 2008, Flin and Patey, 2009)	Organisational Rationalists
Organisational : leadership of teams and organisations	(McFadden et al., 2009, Mohr et al., 2002, Weingart and Page, 2004, Nembhard and Edmondson, 2006)	Organisational Rationalists
Safety: culture, reporting and learning	(Carroll and Quijada, 2004, Ginsburg et al., 2006, Stock et al., 2007, Waring et al., 2007, Weick and Sutcliffe, 2003, Ursprung et al., 2005, Waring, 2005, Vincent, 2004);	Organisational Rationalists
Micro: clinical team view	(Mohr et al., 2004, Brodbeck, 2002)	Professional culture promoters
High risk procedures or event such as medicines management, surgery and anaesthesia, patient falls	(Gauthereau, 2004, Shah and Roberson, 2008, Vincent et al., 2004, Cameron and Kurrle, 2007, Healey et al., 2008, Oliver et al., 2008, Reason, 2005);	Organisational Rationalists and Professional culture promoters
Mortality and Adverse events studies	(Brennan et al., 1991, Chang et al., 2005, de Vries et al., 2008, Leape, 2002, Vincent et al., 2001, West, 2000)	Empiricists
Organisational and capacity issues	Gaba, 2000, Proudlove et al., 2003, Sprivulis et al., 2006, Trzeciak and Rivers, 2003, Richardson, 2006, Weissman, 2007, Hoff et al., 2004)	Organisational Rationalists

Table 2.1 : Overview of main themes in the patient safety literature

A recent review of patient safety research by the World Health Organisation (WHO), identified fifty topics relating to patient safety (Bates et al., 2009). A Delphi process of prioritisation reduced these topics to twenty three major topics which were then categorised further into 'structure'(8), 'process' (5) and 'outcome' (10), although no clear definitions are given for these categories (Jha et al., 2010). It is also recognised that patient safety is often multi-factorial and therefore difficult to understand.

It is identified that developed countries have different priorities from those of developing and transitional countries and that this is reflected in the emergent research areas. The context for this thesis is within developed countries. In this context, the 'lack of communication and co-ordination' followed by 'latent organisational failures' (Bates et al, 2009) are identified as the most important issues to be addressed. Jha et al (2010) explore in more detail the sub-categories of 'structure', 'process' and 'outcome'. The 'structural' factors include the 'breakdown of complex systems' where problems occur at various levels of the system and may involve failures that are not immediately apparent. These can be described as 'latent' or 'hidden' failures. For example, the shortage of staff, long hours of work leading to fatigue, production pressures, poor communication and the poor design of work processes or equipment can all lead to patient safety issues.

From the wide ranging literature it is noted that there is much more to patient safety than just what happens at the interface between patient and clinician (Degos et al., 2009, Walshe and Boaden, 2006, Vincent, 2010, Woods et al., 2010). As Walshe and Boaden (2006) point out, a common theme in the patient safety literature is the 'system or the organisation' contribution to creating error or harm for patients. The common theme within the patient safety literature, that there is a 'system', or 'organisational', as well as a 'human', or 'individual' aspect to safety, is also found in the accident theory literature (Reason, 1990, Reason, 1997, Vincent et al., 1998, Morath and Turnbull, 2005, Leape, 1994, Woods et al., 2010).

It is noted that much of the current literature uses the terms 'system' or 'organisation' interchangeably to convey a variety of concepts with little, if any, underlying theoretical discussion (Table 2.2). This inconsistent use of the terms, combined with the lack of theoretical discussion, means that attributing patient safety failures to system or organisational characteristics has limited explanatory power.

Literature	'System' terms used	Concepts depicted
Reason (1990, 1997)	System or organisational factors	'complex organisations such as nuclear power or air traffic control; control systems; human-system relationship; safety system; complex systems
(Walshe and Boaden, 2006)	System or organisation	Culture; process design; management systems
(Vincent et al., 2004)	Systems approaches	Organisational aspects such as communication, equipment, teamwork, work environment.
(Vincent, 2006)	System model	Human actions have to be understood in 'relation to context'
(Vincent, 2010)	System	Clinical teams as 'micro systems'; hospitals as 'complex adaptive systems'.
(Friesdorf et al., 2007)	System ergonomics	Clinical pathways (macro); standard operating procedures (micro)
(Mohr et al., 2004)	System	'microsystem unit'; 'opaque systems'
(Lowe, 2006)	Systems	Interacting elements (six only)
(Carthey et al., 2001)	System; organisation; institution	A healthcare organisation

Table 2.2: Variety of 'system' concepts found in the literature

It is concluded that the terms system and organisation are widely used in the patient safety literature. However, there appears to be a range of understanding as to what a system in particular means. The term system is often used interchangeably with organisation. There is little theoretical exploration of what constitutes a system or how key concepts from systems thinking might influence the debate about patient safety in healthcare. For example, an important concept about the characteristic of a system is the interaction between the parts and the relationship with the wider environment (Forrester, 1969). Much of the literature does not explore the implication of the dynamic interactions within the system for the safety of patients. There are, however, some models which take an organisational perspective in understanding patient safety, which are considered next.

2.3 System models in the patient safety literature

There are a variety of frameworks in the literature that seek to take account of the organisational factors influencing patient safety. Four in particular are reviewed. These models demonstrate an increasing breadth in setting the boundary around what is meant by 'system' or 'organisation'. The consequence is that more and more 'organisational factors' are included in the frameworks.

Reason (1990) developed the 'Swiss Cheese' model that uses the slices of cheese to convey the idea of a series of barriers which are designed to prevent accidents (Reason, 1997). Like Swiss cheese, the slices have holes in them. When the holes in each slice line up with each other, it represents a series of failures that occur in the defensive barriers, allowing an accident to occur. This model has been widely adopted in the safety literature and applied in healthcare (Reason, 2000, Reason, 2004, Mooney, 2010, Morath and Turnbull, 2005, Wachter, 2008). For example, Wachter (2008) describes the 'Swiss Cheese' model as the widely accepted mental model for 'system safety'.

The complexity of the healthcare 'system' is recognised by Reason (Reason, 2004). However, the theory about the dynamic non-linear nature of hazards and defences is not explored at a conceptual 'systems thinking' level by Reason (1990, 1997, 2000, 2004, 2005, 2008) or by those who draw on his work (Wachter, 2008, Morath and Turnbull, 2005, Mooney, 2010). Instead the Swiss Cheese model appears to rely on the 'direction of causality', or linear model of accident occurrence (Dekker, 2011, Hollnagel, 2004)). The Swiss Cheese model also focuses on the function of the parts rather than the dynamic function of the whole (Hollnagel, 2006). Critical examination of the weaknesses of the 'Swiss Cheese' model recognise that it cannot apply, or be understood, in all situations due to the dynamic nature of accident causation (Reason et al., 2006, Perneger, 2005, Roelen et al., 2010, Hollnagel, 2006, Dekker, 2011).

Vincent et al (1998) make a link between difficult working conditions and the likelihood of making an error in clinical practice. They set out a framework with seven types of 'factors' that impact on patient treatment. These are the institutional context, organisation and management, work environment, team and individual, task factors and patient characteristics. These 'factors' have many sub-factors but the idea of dynamic interactions between all these 'factors' is not explored.

Donabedian (1966) proposed a simple framework of quality assurance in healthcare. It comprises the categories of 'structure', 'process' and 'outcome' (Donabedian, 1966, Donabedian, 2003). He uses the term 'system design' to mean 'structure'. Structure appears to include among many others, the recruitment of staff, the number of hospitals, drug testing, finance and legal protection of participants (Donabedian, 2003). The terms 'process' and 'outcome' are equally wide ranging and inclusive, but the potential interactions between the categories are not explored.

In the Donabedian and Vincent frameworks, which seek to take an 'organisational' view of safety in healthcare, the dynamic interactions between the wide range of 'factors' or 'categories' that they identify are not explored. Therefore, the safety implications of those dynamics are not included, which weakens the explanatory power of both frameworks.

Carayon et al (2006) proposes a 'systems engineering' approach to develop a patient safety framework. He argues that such an approach explains how the 'design of the work system can impact on, not only the safety of patients but also employee and organizational outcomes.' (Carayon et al., 2006). He defines the 'work system' as comprising five components: person; tasks; tools and technologies; physical environment; and organizational conditions which interact and influence each other. Whilst the interactions are recognised and the consequent influences between the components noted, how the interactions influence safety is not fully explored.

Walshe and Boaden (2006) argue that models and frameworks that relate to the organisational aspects of patient safety may not always take account of the 'structures and culture' that are specific to healthcare. They argue that healthcare institutions are unique because of their dominance by the medical profession, the lack of defined processes of treatment, and the close relationship between harm generated by the disease and harm caused by the process of treatment. The argument is made that simply transferring system or organisational models from outside healthcare may not work. This argument is reasonable, but they do not go on and explore the implications for patient safety created by the dynamics of healthcare organisations arising from the unique features they identified.

This review of the literature illustrates that:

- the influence of the ‘system’ or ‘organisation’ on patient safety is widely acknowledged;
- the terms ‘system’ and ‘organisation’ are used inconsistently and interchangeable in relation to patient safety;
- the patient safety frameworks that take a ‘system’ perspective have an increasingly wide scope of ‘factors’ to be taken into account, which assumes a linear and reductionist approach of component failure to explain accidents;
- the exploration of the dynamic interrelationship of the parts of a ‘system’ or ‘organisation’ has not been adequately explored as an influence on patient safety.

Much of the patient safety research takes a reductionist perspective and concentrates on parts of the patient care pathway where many of the systemic issues and interactions are disregarded or marginalised. This can lead to solutions being found for sub-systems, without regard for the interaction with other parts of the wider system. This means that solutions can often be sub-optimal and can produce unintended outcomes.

A key point is that the explanatory power of the assertion of system or organisational influence on patient safety is limited. It is limited because it does not fully explore, at a theoretical level, the dynamic interrelationships that occur between the many parts of any complex system and how they can influence safety. The models and frameworks in the patient safety literature do not fully explore the deeper conceptual issues that arise in complex socio-technical systems (Gilbert et al., 2007, Dekker, 2011). There is strong evidence to suggest that healthcare operates within a dynamic complex socio-technical system (Plsek and Greenhalgh, 2001, Plsek and Wilson, 2001, Braithwaite et al., 2009, Sweeney and Griffiths, 2002).

Given the significance of the characteristics of systems, such as dynamic interactions, on the safety of patients, it can be argued that the literature needs a stronger conceptual basis about ‘systems’ to inform future research and practice. It may be argued that the literature on systems thinking better informs the development of patient safety theory..

2.4 Systems Thinking

This section draws out the underlying principles of systems thinking, which is that the interactions between the parts of a system contribute to the dynamic of the whole (Forrester, 1961). Whilst there is a wide ranging literature on system thinking, it is argued that SD provides the necessary insights to extend the explanatory power of how the ‘system’ influences patient safety. Complexity theory is considered briefly to provide a background to the idea of complex adaptive systems and emergence. SD provides a means to consider the interactions within the system of the ‘elements and functions’ (Hollnagel, 2006) by considering what constitutes the ‘structure’ of the system. The use of ‘causal loop diagrams’ (CLDs) and ‘stock flow diagrams’ (SFDs) are explored as a means to incorporate the constructs of ‘coupling’ and ‘feedback’ that are characteristic of systems. Within SD the interaction of the decision makers is taken into account through the concept of ‘bounded rationality’. That concept seeks to explain the cognitive limitations of humans in the face of complex situations.

2.4.1 What is a system?

The underlying assumption within general systems theory is that the whole is greater than the sum of the parts. The nature of the whole cannot be understood from just studying the parts and that the parts are dynamically interacting (Checkland, 1981, Skyttner, 2005, Forrester, 1961). As noted earlier, when considering safety from a systems perspective it is necessary to look beyond the failure of single components and regard safety as an emergent property of complex systems (Hollnagel et al., 2006).

There are many definitions of systems. Checkland (1981) suggests that a system:

‘...embodies the idea of a set of elements connected together which form a whole, this showing the properties which are properties of the whole, rather than properties of its component parts’ (p.3).

Meadows (2008) defines a ‘system’ as ‘an interconnected set of elements that is coherently organized in a way that achieves something’ (p.11). She argues that a ‘system’ must contain three types of things: ‘*elements, interconnections and a function or purpose*’ (p.11 original emphasis).

The common theme in defining ‘systems’ is the idea that it is the interaction of the parts which creates the whole (Forrester, 1961). What is of interest from a patient safety

perspective is to gain an insight into the non-linear and dynamic nature of the interactions that create the whole. Complex socio-technical systems have multiple interactions leading to different levels of linking and feedback between the parts (Sterman, 1994, Braithwaite et al., 2009, Carayon, 2006, Reiman and Oedewald, 2007, Lane, 2001a). Another key feature of systems found in healthcare is that they are opened to their environment. Therefore, part of the contribution to the behaviour of the whole is the interaction with the context within which a system operates (Cummings et al., 2001).

Another way of describing systems is as ‘complex adaptive systems’ (Holling, 2001). The basis of much of the complexity science is that of an individual agent. Complexity theory studies the behaviour at the micro or individual level as a way to explain the working of the wider system. It is the interaction of the agents that create the emergence and the coherence that might be described as a system. Often a set of simple rules generate the behaviour. For example, flocking birds keep a certain distance from other birds and fly in the same direction (Lewin and Regine, 1999). It is argued that underlying any complex system there are a few simple controlling rules or processes (Holling, 2001).

Chaos and complexity theory provide a number of helpful metaphors that can apply to organisations and systems (Lewin and Regine, 1999). Using the metaphors provides a means to recognise systems less as mechanistic constructs and more like living entities where relationships, attractors and fuzzy boundaries are at the core of understanding the dynamics (Lewin and Regine, 1999, Plsek and Greenhalgh, 2001). Complexity science has been applied to a variety of healthcare settings in an effort to gain an understanding of the adaptive nature of people and organisations (Sweeney, 2006, Plsek and Greenhalgh, 2001, Plsek and Wilson, 2001, Sweeney and Griffiths, 2002, Kernick, 2004, Sweeney and Williams, 2010). The insights from complexity theory about the nature of systems being more like living entities rather than mechanical, is regarded as helpful in understanding adaptive and emergent behaviour.

Complex systems have emergent properties (Hollnagel, 2004). This means that the relationship between the parts in a system create properties, which are more than just the sum of the parts (Dekker, 2011). Therefore, to understand safety by taking a linear or reductionist approach is limited to providing a component based answer as to why

things go wrong. A systems thinking approach to complex systems considers the interaction of the parts and their emergent properties in providing an explanation of the behaviour of the whole system. As Dekker (2011), argues, a systems approach to complex systems is to 'go up and out, not down and in'.

There are some key characteristics of complex systems which differentiate them from complicated ones (Dekker, 2011, Cilliers, 1998). These are summarised as being:

- Complex systems have a large number of parts which interact physically or through the passing of information.
- The interactions are non-linear and produce direct or indirect feedback onto the interaction activity. Such feedback can introduce a multiplier effect.
- Complex systems are open in that they interact with and influence their environment. Establishing the boundary of any complex system is therefore a matter of judgement.
- Within complex systems each part is unaware of the behaviour of the system as a whole and the full consequences of its actions. Parts respond to locally available information. Complexity arises from the patterns of interaction between the many parts.
- Complex systems are a product of their history. The dimension of time is important in providing a deeper understanding of the diachronic processes at work.
- Complex systems operate under conditions where there is a constant flow of energy. Systems that experience equilibrium do not survive.

A system thinking approach can therefore provide a means to develop the understanding about the dynamic processes within complex socio-technical systems, which potentially impact on patient safety. Complexity theory does provide some helpful metaphors, but is limited in explaining the detail of the interactions. To help us consider the interactions of a complex socio-technical system, the contribution from the SD literature is considered in the next section.

2.4.2 System Dynamics

SD starts with the assumption that systems are complex with multiple interactions.

Rather than regarding complexity as being unknowable, SD claims to be able to model

the complexity to a greater extent than other system approaches (Rodriguez-Ulloa and Paucar-Caceres, 2005). Jackson (2003) places SD in the functionalist category in his assessment of system thinking methods. However, there are others who argue that SD has built bridges towards a more interpretive and pluralist paradigm (Lane, 2001b, Rodriguez-Ulloa and Paucar-Caceres, 2005, Lane, 2000a, Senge, 1990, Lane, 2001a). SD is therefore linked to social theory that takes a dialectical approach to agency and structure. That means the decision making agents are not passive in the presence of the structures they work within. Rather, there is a dynamic interrelationship between the agent and the structure. The agent is influenced by the structure and, in turn, the agent influences the structure (Lane, 1999). It is important to note that SD focuses on the aggregate pattern of behaviour rather than the individual actions of agents. SD treats 'causes as pressures which produce aggregate patterns of behaviour' (Lane, 2000a).

SD takes the view that the 'structure' is a key factor on the overall behaviour of the system (Jackson, 2003). There are many variables in the 'structure' that interact. The interactions within the system itself contribute to the dynamics, which are either sustaining or destructive, through either positive or negative feedback (Forrester, 1958, Forrester, 1961, Forrester, 1969). These ideas have been developed and are used in modelling a wide range of complex systems (Sterman, 2000, Sterman, 2001, Morecroft, 2007, Meadows, 2008, Senge, 1993). The underlying theory is that non-linear dynamics and feedback control can be applied to the behaviour of physical, technical and human systems (Sterman, 2000). Sterman (2000) argues that the behaviour of the system is derived from the interaction of the structure of a system with the human decision making process:

'The behaviour of a system arises from its structure. That structure consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it.' (Sterman, 2000, p.107)

Meadows (2008) describes the 'structure' of a system to consist of 'the interlocking stocks, flows and feedback loops' (p.89). It is the structure that determines the hidden behaviours in the system. For Meadows (2008) the behaviours appear over time and can be analysed using time graphs. Meadows (2008) suggests that when taking a systems thinking approach, it is trying to understand the relationship between an 'event', the resulting 'behaviour' (e.g. oscillation) and the 'structure' of the system.

Sterman (2000) argues that it is the 'stocks' that depict the character of the system and provide the information on which many of the decisions are taken. Stocks collect the effects of past events and decisions. 'Stocks' are the accumulations in the system; usually derived from the difference between the 'flow' into the process and 'flow' out. Differences in flow rates can create backlogs or scarcity in a system. For example, the 'stock' of greenhouse gases in the atmosphere is a result of previous production of those gases (inflow). They can only be reduced by the rate of outflow being greater than the rate of inflow.

A key feature of 'stocks' is that they can 'decouple rates of flow and create disequilibrium dynamics' (Sterman, 2000, p.196). In effect the stocks act as a buffer to the differences in flow rates. Each rate of flow may be governed by different decision making processes. The state of the stock acting as a buffer is part of the feedback loop mechanism (Sterman, 2000). In applying this method to hospitals the 'flow' relates to patients being admitted treated and discharged, and the 'stocks' are the wards or departments.

SD studies have two ways to illustrate and model system behaviour. These are SFDs and CLDs (Lane, 2008). SFDs can be used to illustrate the flow of work through a system. The notion of a system boundary is important when modelling a system using feedback loops, stocks and flows. When building a model, the decision of where to set the boundary is a judgement depending on the purpose of the analysis (Midgley, 2000, Meadows, 2008). When depicting a model using an SFD or CLD there are a number of conventions that have evolved (Lane, 2008) which are presented in Appendix 2.1.

Senge (1990) and Meadows (2008) set out a series of system archetypes which suggest common dynamics that occur in systems. The identified archetypes provide potential levers for changing the dynamics. Sterman (2000) similarly sets out 'modes of dynamic behaviour' ranging from 'exponential growth', 'oscillation' to 'overshoot and collapse'. Oscillation occurs when the performance of the system fluctuates considerably. Sterman (2000) points to the idea of 'local stability' where perturbation on a system will cause it to oscillate and then return to the same point of equilibrium. However, when there is 'local instability' small disturbances can move the system further away from the original point of equilibrium. Sterman (2000) illustrates his point suggesting that a ball

balanced on top of a hill is in local unstable equilibrium. A slight breeze pushes the ball off the top of the hill where positive feedback occurs and the ball accelerates down the hill until eventually coming to rest at the bottom in a new state of equilibrium. There are therefore different states of equilibrium to be considered.

CLDs can depict the influences and the linkage, known as 'coupling', between the parts of a system. CLDs also show whether the system is in balance or whether there is a reinforcing loop dominating, which causes the system to lose equilibrium. CLDs can therefore depict the changes in the 'coupling' of the parts and the overall impact this has on the whole system

Concluding this review of SD the following points are noted:

- the origins of SD are in the hard and determinist paradigm. However, bridges have been built towards combining human agent and structural concepts that allows SD to contribute to interpretive social theory (Lane, 2000a);
- the working definition of a system is that it '...embodies the idea of a set of elements connected together which form a whole, this showing the properties which are properties of the whole, rather than properties of its component parts' (Checkland, 1981)
- Meadows (2008) argues that a system thinking approach seeks to understand the relationship between an 'event', the resulting 'behaviour' (e.g. oscillation) and the 'structure' of the system;
- the 'structure' of a system '...consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it.' (Sterman, 2000) This definition brings together the structural and human agent aspect of complex systems;
- two diagrammatic methods are used to model the dynamics of systems; SFDs and CLDs;
- a number of system archetypes have been identified in the literature, which suggests points of intervention to change the behaviour;
- there can be different points of dynamic equilibrium in a system;
- SD has limitations based on the ability of the user to understand the problem situation, conceptualise and build the model and interpret the results.

It is argued that concepts derived from system thinking literature, and SD in particular, can assist in developing the understanding about how the 'system' influences patient safety. A key feature of healthcare is the human agents. Sterman (2000) argues that it is possible to include decision makers in SD modelling. Therefore, the final area to consider is the concepts used by SD to depict the decision making agents and process within a system.

2.4.3 Decision Making

As noted in SD, the 'structure' of the system is defined as being the stocks, flows and feedback interacting with the decision making agents. SD seeks to take account of the interaction with decision makers in the modelling through the concept of 'bounded rationality' (Sterman, 2000).

The nature of human cognitive capacity is that it is soon overwhelmed by complex problems and is unable to make objectively rational decisions (Simon, 1957). The limited ability to process all the information arising, particularly in dynamic complex situations, is known as the concept of 'bounded rationality' (Simon, 1982, Morecroft, 1983, Sterman, 1989, Meadows, 2008). This means that there are clear limits to the rational capabilities of decision makers.

Decision makers adopt a number of techniques to manage the cognitive workload in dynamic complex situations (Sterman, 2000). These include habit and routines where there is little cognitive effort or simplified rules for making decisions, such as check lists (Loukopoulos et al., 2009, Gawande, 2010). Another technique is setting specific goals or targets which then focus the attention of decision makers towards certain information with the consequence of potentially ignoring other issues (Crilly and Le Grand, 2004, Barber, 2008). A similar method is to sub-divide the problem into manageable parts. The parts are passed down the line, within clear parameters, with the idea that the cumulative effort of many decision makers will solve the problem. Decisions are often made on the assumption that there are no time delays, undesired consequences, feedbacks or non-linear interactions. This behaviour has been termed 'intended rationality' (Sterman, 2000) and is characteristic of a reductionist approach to problem solving.

Meadows (2008) describes how people make rational decisions based on the information they have. Decision makers often have limited information. For example, those taking decisions in complex systems are not always aware of what other decisions are being made at the same time. Meadows (2008) suggests that:

‘We misperceive risk, assuming that some things are more dangerous than they really are and others much less. We live in an exaggerated present – we pay too much attention to recent experience and too little attention to the past, focusing on current events rather than long-term behaviours.’ (p.107)

Meadows (2008) argues that ‘bounded rationality’ is often caused by system archetype structures as people fail to see the wider SD which leads to unforeseen and unhelpful consequences. Bounded rationality influences the world view, or mental model, that we hold which affects our approach and decision making (Kuhn, 1970, Senge, 1990, Argyris, 1999, Doyle and Ford, 1998). The literature provides a range of insights in this area. Snowden and Boone (2007) see the world as having three types of systems: ordered (simple and complicated), chaotic and complex. They developed this into what is known as the Cynefin Model. The agents in an ordered system are constrained by rules and procedures. Within a chaotic system agents are unconstrained and act independently of each other. In a complex system agents are lightly constrained, they interact with each other and modify the system (Snowden and Boone, 2007).

The inclusion of decision making agents in SD is an important feature in seeking to understand the behaviour of systems. The concept of ‘bounded rationality’ depicts the limitations of human decision makers and their inability to comprehend all the information and feedback that occurs in complex situations. Humans have a number of techniques and predetermined world views that help them make sense of their environment. These techniques and mind-sets can lead to decisions being taken without due regard to feedback, delay and unintended consequences that are common in complex socio-technical systems. Capturing the decision making process in an SD model is not easy. However, in modelling a system it may assist decision makers in seeing where and how their interventions influence the behaviour of the whole as well as the parts of the system.

2.4.4 Summary of systems thinking contribution

The underlying principles of systems thinking is that the interactions between the parts of a system contribute to the dynamic of the whole (Forrester, 1961). An SD approach

to understanding systems focuses on the dynamics that occur between the feedback loops, stock and flows interacting with the decision making processes.

SD provides insights into the dynamics interactions within systems by taking account of:

- the flow of work;
- the accumulations of work (stocks);
- the information and decision making that influences the stock and flows;
- the feedback and coupling between the parts of the system;
- the equilibrium of the feedback dynamics for the system.

The features identified above are found in complex healthcare systems. The argument from the patient safety literature is that ‘system’ factors are regarded as a major contributor to failures. It is argued that the insights from SD about the key characteristics help to develop the explanatory power of how the ‘system’ influences patient safety.

However, SD has only been used to study the operational management of hospitals (Lane et al., 2000, Lane and Husemann, 2008b). It has not been applied previously to consider the implications of the system behaviour for patient safety. Other approaches have considered the influence on safety of the systemic features of organisations, such as ‘feedback’ and ‘coupling’, which are reviewed in the next chapter.

2.5 Summary

The extant literature takes account of ‘system’ or ‘organisational’ factors that influence patient safety. However, the underlying theory to explain the influence of system factors is weak. Organisational and systems models used in the literature have been reviewed. Weaknesses in the explanatory power of those models are evident as there is little exploration of the dynamic interactions between the parts that creates the behaviour of the whole. Those interactions are characteristic of complex systems. It is argued that the weakness can be addressed by developing the conceptual framework about patient safety from the systems thinking literature.

The definition of a system used is from Checkland (1981), who argues that a system:

‘...embodies the idea of a set of elements connected together which form a whole, this showing the properties which are properties of the whole, rather than properties of its component parts’

The nature of healthcare organisations is that they operate in complex socio-technical environments (Plsek and Wilson, 2001). Therefore, the characteristics of such environments and the implications for systems have to be included when considering patient safety. There are some key points derived from the review of the systems literature which have application when considering the safety of patients:

- complex systems have a large number of parts which interact and are open to their environment. The interactions create feedback which influences the behaviour of the system as a whole;
- social systems include human agents as decision makers who interact with other agents, the technical and work flow parts of the system and the environment. The interactions are dynamic;
- the ‘structure’ of a system ‘...consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it.’ (Sterman, 2000, p.107) This definition brings together the structural and human agent aspect of complex systems;
- insights can be gained from using systems thinking into the relationship between an event, the resultant system behaviour and the ‘structure’.

The extant literature is limited in taking account of these key points when considering the system factors that influence patient safety. In the next chapter the wider safety, or accident theory, literature is reviewed to establish how concepts from systems thinking are incorporated and how they can help inform the patient safety literature.

Chapter 3 – Developing a system resilience approach

3.1 Introduction

The extant patient safety literature makes the clear point that the ‘system’ creates risk for patients. However, the system based models used lack explanatory power. The main reason for that weakness is the failure to synthesis concepts from the systems thinking literature. The wider safety literature does make some use of systemic concepts. This chapter reviews the accident theory literature before examining the concept of resilience. Whilst resilience has been developed for use in a number of academic disciplines, it is increasingly being considered to provide insights into safe practice in complex systems. The resilience engineering literature has developed this approach and in particular the model of a SWE. The model is conceived as a means to depict a system operating within the constraints of the envelope boundaries whilst subject to competing pressures. A synthesis of concepts from accident theory with the SWE model is suggested. There are some weaknesses with the model in being able to capture all of the key characteristics of systems identified in Chapter 2. However, it is argued that using such a model will assist in expanding the explanatory power of how the ‘system’ influences patient safety.

3.2 Concepts from Accident Theory

There are a number of key concepts from accident theory that can contribute to informing a systems level understanding of patient safety. The literature is wide ranging, with a variety of academic disciplines contributing. The field is more mature than patient safety and has often developed theory from the detailed study of major accidents or disasters (Dekker, 2011). This review is not exhaustive, but seeks to identify those concepts which help to inform a systems based theoretical framework to improve the understanding of how the ‘system’ influences patient safety.

Reason argues that there are three aspects to accidents: ‘universals’, ‘conditions’, and ‘causes’. ‘Universals’ describe the tension between production and safety that is frequently found in organisations. Within the literature on error in organisations there is a wide recognition of the often conflicting priorities of production and safety (Flin et al., 2008, Reason, 1997, Dekker, 2005). Reason (1997) argues that, on the one hand, to

focus on safety at the expense of production is the route to bankruptcy. On the other hand, to concentrate on production without adequate attention to safety will lead to catastrophe. The methods of production in an organisation are usually well understood, measured and managed. The safety functions, however, tend to be less well defined and measured (Reason, 1997). A consequence of this difference, argues Reason (1997), is a tendency to for managers to favour production goals over safety.

Woods (2006) describes the production requirement as an ‘acute goal’ which has to be balanced with the ‘chronic goal’ of safety. A factor underlying this bias is that safety is a ‘dynamic non-event’ (Weick, 1987). What this means is that there is work constantly being done to maintain safe operations. The outcome is that no accidents happen. Due to the lack of safety incidents, managers assume that their decisions, for example, to increase production, will not impact upon safety. What is often not realised by decision makers is the continuous and sometimes increasing effort made by staff to achieve ongoing safety.

Reason (2008) argues that it is the ‘universal’, (the favouring of production over safety), which create the second aspect of accidents - ‘conditions’. These ‘conditions’ have also been described as ‘latent factors’ which weaken the defences against safety failures. It is argued that the ‘conditions’ create the situation where the third aspect of accidents can occur - ‘causes’. An accident can be ‘caused’ when various factors combine to breach the safety defences. This is illustrated by the holes lining up in the slices of the Swiss Cheese model (Reason, 2008).

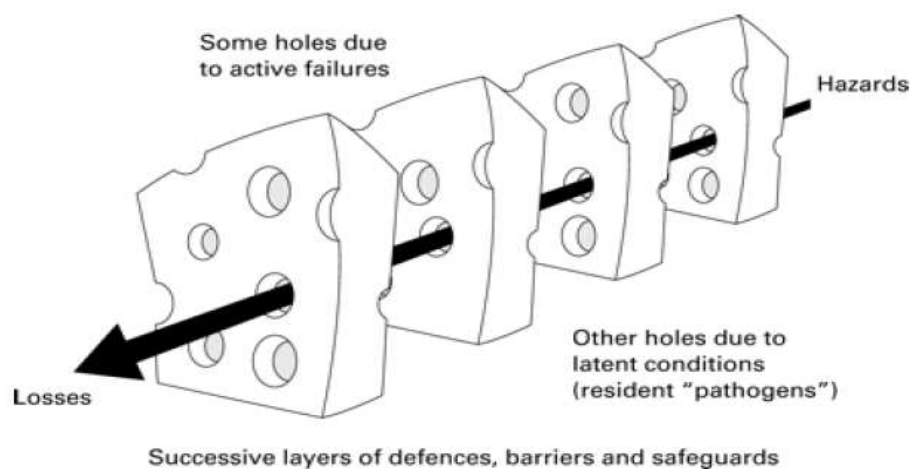


Figure 3.1: Reason's Swiss Cheese model (Reason, 1997)

Reason (1997) suggests that to achieve improved safety, attention has to be paid to the sharp and blunt ends of the system (Figure 3.2). Those at the sharp end are those people delivering patient treatment. They are human and are likely to make mistakes (Reason, 1990). Equally those managers with responsibility for the organisation make decisions at the blunt end (removed from direct delivery), which can create ‘conditions’ at the sharp end that are more susceptible to error. The decision, for example, to increase production, can create ‘conditions’, that combine with local workplace factors at the sharp end making staff more likely to make an error.

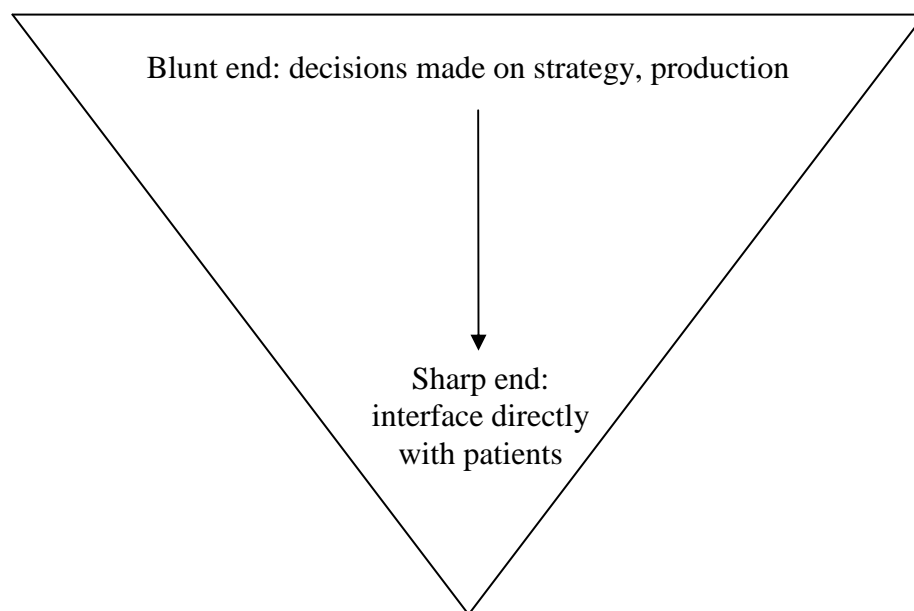


Figure 3.2: Blunt and sharp end of an organisation
(adapted from Woods et al, 2010)

It is argued that Reason (1990, 1997) takes a fundamentally linear approach with the Swiss Cheese model. The same criticism can be made to the sharp and blunt end concept, which appears to simplify an ongoing dynamic relationship.

In further developing the conceptual basis for patient safety from a ‘systems’ perspective it is necessary to use concepts that emphasise the interactions of the parts of a system (Forrester, 1969), noted in Chapter 2. Within the accident theory literature there is a growing body of work that seeks to build a theoretical framework to help explain the influence of the ‘system’ on safety. Researchers who use a systemic model

to understand the 'functional characteristics of the system' (Hollnagel, 2004) use the concepts of 'coupling' and 'feedback' between the parts of a system.

The concept of 'coupling' is used to describe how closely linked together parts of a system are (Cook and Rasmussen, 2005, Perrow, 1984, Perrow, 1999, Paté-Cornell, 1993, Beekun and Glick, 2001, Rijpma, 2001). Systems that are loosely coupled have different departments with very little, if any, interaction between them. Where there is tight coupling the parts of a system interact extensively. In Normal Accident Theory (NAT), Perrow (1984, 1999) argues that where complex technological systems are interactively tightly coupled, accidents are an inevitable and should therefore be regarded as 'normal'. One of his cited examples of an interactive tightly coupled system is the Three Mile Island Nuclear Power Plant accident. Perrow (1984) describes such failures as 'normal' or 'system' accidents. He suggests that such accidents should be expected, and there is little that can be done to reduce their likelihood. He does however, make suggestions to reduce the vulnerability of such complex systems (Perrow, 1999, Perrow, 2007). The point made from a safety perspective, is that where there are more dynamic interactions between the different parts of the system, the risk of an accident increases. This increased risk is due to the dynamic and sometimes unpredictable 'feedback' that occurs between the parts creating unforeseen conditions that can cause accidents (Diehl and Sterman, 1995, Meadows, 2008, Sterman, 1989, Lane and Husemann, 2008a, Senge, 1990).

'Feedback' is where an interaction between the parts creates a situation which influences future interactions of the same or other parts in a system (Sterman, 2000). There are two types of feedback: 'reinforcing' and 'balancing'. These two types of feedback will be explored in more detail below. Where the systems are complex, with many feedback loops combined with tight coupling of the parts, then the risks of a system accident are much higher (Perrow, 1984).

As noted earlier, Walshe and Boaden (2006) argue that healthcare is a system made up of many parts with both social and technical complexity. The concepts of 'coupling' and 'feedback' will therefore help to inform the conceptual basis for patient safety at a systemic level. Other studies, taking a High Reliability or Normal Accident approach, have sought to understand how complex organisations have reduced the risks to safety that result from tight coupling and multiple feedback (La Porte, 1996, Roberts and Bea,

2001, Schulman et al., 2004, Weick, 1987, Weick and Sutcliffe, 2001, Perrow, 1994, Rijpma, 2001, Sagan, 1993).

One key aspect arising from High Reliability Theory (HRT) and NAT is the use of 'redundancy'. Redundancy can also be thought of as 'buffer capacity' where resources (people and equipment) are available, if needed, to prevent unexpected failure. Resources are termed 'redundant' because they are there just in case they are required, and have no other function. In HRT studies, such as those conducted on naval aircraft carriers, it was found that there is a high use of 'redundancy' in the form of people, equipment and processes to ensure that any error or failure by one part, is caught and corrected by an otherwise unused (redundant) resource (Roberts, 1990). Conversely, NAT theorists argue that redundancy can add to the complexity of the system design and potentially lead to complacency in an organisation (Sagan, 2004). Complacency arises when people in one part believe, that with multiple layers of protection against failure, some other part will prevent anything going wrong, so their vigilance drops. It is argued that the use of redundancy is ineffective in reducing accidents (Sagan, 2004). Both HRT and NAT views about the use of redundancy have their validity, and need to be considered in the context of specific organisational design and operation.

The ethnographic study of the Space Shuttle Challenger mission that exploded on take-off, proposed the 'normalisation of deviance' concept (Vaughan, 1996). Vaughan (1996) explains how, within a complex technically orientated organisation, processes were established that recategorised the degree of risk to one which was deemed acceptable. The revised view of risk therefore became the 'normal' in such a situation. Studies in healthcare have found 'normalisation' by staff to inpatient falls in hospital (Williams et al., 2010) and errors in operating theatre in hospitals (Waring, 2005, Waring et al., 2007). The process of normalisation results in low levels of reporting of incidents, as staff regard error or harm to patients as part of the process of treatment. For example, high mortality rates in certain surgical procedures are explained as being due to treating more 'difficult' cases (Boseley et al., 2010). The apparent high death rates for emergency patients in a particular hospital being explained initially as a data collection problem and not reflecting reality (Healthcare Commission, 2009a).

The research into the shooting down of two American Black Hawk helicopters in a 'friendly fire' incident in the no-fly zone in northern Iraq (Snook, 2000) provides an

insight into why written procedures, designed to ensure safety, are deviated from over time. From his analysis, Snook (2000) developed the concept of 'practical drift' to explain how people in the work place slowly, over time, departed from the written procedures. The 'drift' from written procedures is driven by practical necessity. The 'drift' occurs over a period of time and without adverse consequences. The point is reached when tight coupling is rapidly re-established due to unforeseen events, and a failure comes as a surprise. This concept of 'drift' also applies to the process of normalisation (Dekker, 2011). As Vaughan (1996) argues, the process of recategorising risk happens slowly over time. When nothing disastrous goes wrong, the new category of safe becomes the norm. This process can then be repeated with the consequent 'drift' towards unsafe practice. The movement towards unsafe working has also been termed violations and migrations from the standard procedures (Amalberti et al., 2006).

Further insights into systemic concepts that can underpin patient safety are available from the literature that takes a 'human factors' perspective when looking at complex organisations (Carayon, 2006, Mooney, 2010, Reason, 1995, Woods et al., 2007, Dekker, 2006). There are helpful points for us to consider, which include the multiple interfaces between people, technology, and goals. As Dekker (2006) argues, safety is not the only goal and has to be achieved when staff are pressured, in situations of ambiguity and often with unreliable technology. People in dynamic complex socio-technical systems are often in a position where 'trade-offs' decisions have to be made between doing their job efficiently or thoroughly (Hollnagel, 2009a, Woods et al., 2010).

In summary, there are a number of concepts used in the accident theory literature which are of interest when thinking about patient safety at a system level:

- production verses safety
- blunt / sharp end
- latent or hidden conditions
- safety as a dynamic non-event
- coupling and feedback of interacting parts of a system
- redundancy / buffer capacity
- normalisation
- practical drift (migration)

- trade-offs

These concepts help to inform a ‘system’ based understanding of patient safety where the dynamic interaction of the parts is taken into account. A recent development in the literature, that seeks to encompass many of these concepts, is the idea of ‘resilience’. Resilience itself is a broad concept which is beginning to be used in the safety literature and has been applied in a limited way to healthcare (Carthey et al., 2001, Jeffcott et al., 2009). The next section explores ‘resilience’ as a means to provide a unifying model for many of the identified concepts from accident theory and systems thinking. The aim is to identify a model that takes account of the conceptual issues underlying the dynamic interrelationship of the parts in the systemic approach to better understand patient safety.

3.3 Resilience

‘Resilience’ is increasingly being used in the literature in relation to safety (Smith and Fischbacher, 2009, Sheridan, 2008). The aim of this section is to review the literature on resilience and to identify the main themes that take account of dynamic interactions between the parts of a system. ‘Resilience’ as a concept is defined as being both proactive and reactive. It is regarded as part of a continuum that includes the concept of vulnerability. To take account of the dynamic and socio-technical nature of healthcare systems, the approach of resilience engineering is used, and in particular, the model of a SWE within which a system operates. It is argued that the SWE model can unify many of the concepts that are identified from the accident theory literature. However, the model does have limitations that are identified.

The concept of ‘resilience’ is studied in a number of academic disciplines including: child development (Luthar et al., 2000), disaster management (Somers, 2009) and ecology (Cropp and Gabric, 2002, Adger, 2000, Holling, 1973). Each field defines ‘resilience’ differently and therefore undertakes research from different perspectives (Manyena, 2006). Much of the literature defines ‘resilience’ as a characteristic that allows for recovery from a perturbation, such as a natural disaster or an abused childhood (Berkes, 2007, Boin and McConnell, 2007, Cropp and Gabric, 2002). Masten (2007), writing in the child development literature, regards ‘resilience’ as referring to the ‘capacity of dynamic systems to withstand or recover from significant disturbances’.

She suggest that such a concept can be studied at different levels; from the global to micro and over different time periods (Masten, 2007).

A contrasting view is that ‘resilience’ refers to the ability to absorb a disturbance, rather than just recover from one (Adger, 2000) (see Figure 3.3). Woods (2006) defines resilience in relation to patient safety as ‘a work system’s ability to buffer, adapt to, absorb and prevent adverse patient outcomes in the face of disruption.’ Again this proposes the idea of ‘resilience’ being more than just recovering from a perturbation. Similarly, Hollnagel defines a resilient system as being able ‘to continue to perform as required’ before during or after a disruption or continuous stress (Hollnagel, 2009b).

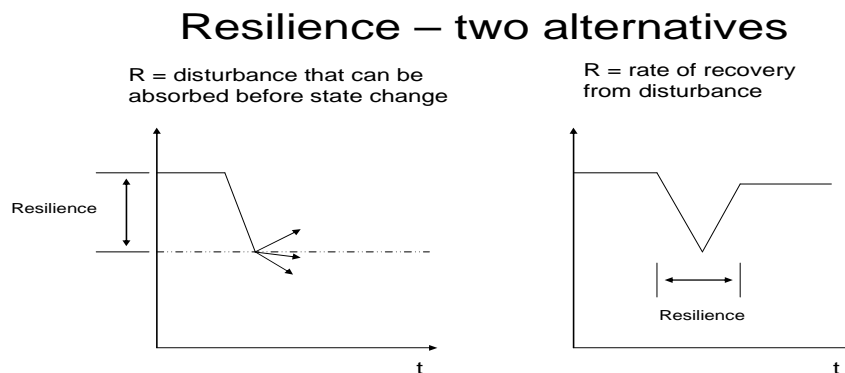


Figure 3.3: Two views of ‘resilience’
(adapted from Adger (2000))

It is argued that it is possible to view ‘resilience’ as being both / and, rather than one or the other. In other words, ‘resilience’ can refer to the ability to absorb or adapt to a disturbance, as well as the more traditional idea of recovering after a perturbation. There are many definitions of ‘resilience’ in the literature (Amin and Horowitz, 2008, Weick and Sutcliffe, 2001, Luthar, 2003, Holling, 1973, Sheridan, 2008, US Government Accountability Office, 2009, Dalziell and McManus, 2004).

The definition of resilience used in this research is:

‘...the ability of an organisation (system) to keep, or recover quickly to, a stable state, allowing it to continue operations during and after a major mishap or in the presence of continuous significant stresses.’ (Wreathrall, 2006)

This definition uses the idea of ‘a stable state’ as being a core requirement for an organisation or system to function. It also includes the view that resilience is about the ability to absorb a disturbance. By anticipating a disturbance, it becomes easier to maintain a stable state during a disruption or stress, as well as to recover from such an event. The definition allows for both types of resilience identified by Adger (2000). The Wreathrall (2006) definition also allows for not just perturbations, but also ‘continuous significant stresses’. Placing an organisation or system under continuous pressure can create ‘latent conditions’ and cause a ‘drift to danger’ (Rasmussen, 1997), or ‘practical drift’ (Snook, 2000).

‘Resilience’ as a concept looks at the characteristic of a system in terms of its ability to withstand disruption or stress. It may also help to consider the characteristic of a system in terms of the inability to adapt and continue in a stable state. The concept of ‘vulnerability’ in parts of the literature is associated with ‘resilience’; often being regarded as the different sides of the same coin (Vogel et al., 2007, Dalziell and McManus, 2004, Luthar, 2003). However, both ‘resilience’ and ‘vulnerability’ are concepts with multiple definitions (Manyena, 2006).

Writing from the child development field, Luthar and Zelazo (2003) debate, without arriving at a clear conclusion, whether the constructs of ‘protection’ (resilience) and ‘vulnerability’ are on the same continuum or whether they are distinct entities (Luthar and Zelazo, 2003). Manyena (2006), in considering ‘resilience’ and ‘vulnerability’ in the context of disaster management, argues that they are separate entities rather like job satisfaction and job dissatisfaction. ‘The absence of job dissatisfaction does not mean that you have job satisfaction’ (Manyena, 2006 p.443). However, it is suggested that the logic of this argument, that there are separate entities, can apply equally the other way. The presence of job dissatisfaction does mean that you will *not* have job satisfaction; where there is job satisfaction there will *not* be job dissatisfaction. So a logical case can be made, (with suitable definitions) that where there is ‘vulnerability’ to perturbations then ‘resilience’ is weak or not present. Where there is ‘resilience’ to perturbations then the ‘vulnerabilities’ will be minimal or non-existent. Manyena (2006) does concede that when ‘vulnerability’ is defined in a way which relates to ‘the degree of capacity, then vulnerability is closely associated with the level of resilience’ (Manyena, 2006 p.440). The position adopted within the thesis is that ‘vulnerability’ and ‘resilience’ are on the same continuum.

The definition of 'vulnerability' used is:

'Vulnerability is the degree to which a system acts adversely to the occurrence of a hazardous event. The degree and quality of the adverse reaction are conditioned by a system's resilience (a measure of the system's capacity to absorb and recover from the event).' (Timmerman, 1981 quoted in Manyena, 2006, p.441.)

Most definitions only deal with vulnerability at a component level. This definition takes a higher level perspective, which is consistent with this thesis. It also makes the link explicit between the nature of the adverse reaction and the 'resilience' of the system to both absorb and recover from a perturbation.

Writing about resilience from an ecological perspective, Holling (2001) synthesises the idea of hierarchies and adaptive cycles to create the theory of 'panarchy'. Hierarchies are used to describe the 'semi-autonomous levels' of adaptive cycles that interact rather than a method of top-down control. The lowest level is the individual; the highest is the political and cultural. The higher levels adapt at a slower speed than the lower levels. The model depicts a dynamic hierarchy where the levels pass on information or material to enhance the ability to adapt in the face of change. The term 'panarchy' is used to overcome the inflexible associations with the word 'hierarchy'. Panarchy is used as a 'representation of a hierarchy as a nested set of adaptive cycles' (Holling, 2001).

The key point to take from Holling (2001) is the idea of different levels of adaptive cycles that have a dynamic interrelationship. When experiencing perturbation, each level may exploit their properties of adaptive capacity in different ways, which may either compliment or conflict with other levels.

The main points to draw from this section on resilience are:

- 'resilience' is a concept studied by many academic disciplines with different definitions;
- 'resilience' as defined for this thesis relates to the ability of the system to both absorb and recover from perturbations or significant continuous stress;
- 'resilience' in relation to patient safety is about the ability to prevent harm to patients when the system experiences perturbations or significant continuous stress;

- ‘resilience’ is unlikely to be present where there are aspects of the system which are ‘vulnerable’ to losing their stable state in the face of perturbations or significant continuous stress;
- ‘resilience’ can be modelled as ‘panarchy’ - a nested set of adaptive cycles.

The concepts of ‘resilience’ and ‘vulnerability’ have been developed in relation to concepts about ‘systems’ in the literature on ecology (Cropp and Gabric, 2002) and more interestingly, for the purpose of this thesis, in ‘resilience engineering’, which are considered next.

3.3.1 Resilience Engineering and the Safe Working Envelope

‘Resilience engineering’ explores a systemic perspective of ‘resilience’. It includes the ability to anticipate, absorb, mitigate and recover from a disturbance (Woods, 2006). Hollnagel (2006) argues that to understand accidents it is necessary to look at the dynamic concurrences which take place ‘amongst events and functions rather than among components’ (Hollnagel, 2006). This means looking beyond the failure of single components in a system and regarding safety as an emergent property of complex systems (Hollnagel et al., 2006). The challenge is to find out why a system may lose its dynamic stability and become unstable. Woods (2006) suggests that ‘resilience engineering’ is an approach to safety that ‘focuses on how to help people cope with complexity under pressure to achieve success.’ It is not about just counting error and then acting to reduce that count, but rather putting safety as a core value in an organisation (Woods, 2006).

Resilience is a proactive approach to safety which, if successful, means that the system can keep operating in what resilience engineering theorists describe as an ‘operating envelope’, within which the system is designed to be competent (Rasmussen, 1997; Woods, 2006). The SWE is a conceptual model based on assessments of uncertainty and the ability of the system to manage the resulting change. Resilience is about ‘monitoring the boundary conditions of the current model of competence’ (Woods, 2006). The aim is to assess the adaptive capability of the system to remain within the safe operating envelope. The ability of the system to remain within the envelope is described as ‘system resilience’.

There are a number of models of ‘system resilience’ (Woods et al., 2009). These include the ‘ball and cup’ metaphor and the ‘stress-strain state space’ (Woods and Wreathrall, 2008, Woods et al., 2009). The latter model is used in material science and is applicable to those situations where a material, individual or organisational capacity is stretched to a point beyond which it cannot recover. Whilst this model does provide some insights into resilience properties, such as critical breaking points, it is limited in taking account of the complex socio-technical aspects of healthcare. The ball and cup model provides insights into the amount of disturbance a system can survive. Again this model appears limited in taking account of the interactions and competing pressures found in complex systems and it is not widely used in the resilience engineering literature. It may therefore be argued that these models are underdeveloped in comparison with the work of Rasmussen (1997). Rasmussen’s (1997) use of the safe working envelope (SWE) model takes account of the wider social context as well as the competing dynamic interactions. Safety is influenced by the wider context within which a system operates, from government, through regulators and organisations to workers (Rasmussen, 1997). This suggests that safety is embedded within a complex adaptive socio-technical system, with many stakeholders and interactions. As such, safety is more complicated than just setting out rules to be followed and behaviour to be controlled.

Rasmussen (1997) is widely cited in the literature on safety, risk and control theory across industries ranging from shipping to information technology. Some of the safety literature relating to healthcare cites Rasmussen (1997), for example, (Miller and Xiao, 2007, Jeffcott et al., 2009a, Jeffcott et al., 2009b, Armitage, 2009, Dekker et al., 2011, Wears, 2005, Wears, 2004, Wears and Cook, 2010, Wears et al., 2008, Amalberti et al., 2006, Sittig and Singh, 2010, Karsh and Brown, 2010, Cook and Rasmussen, 2005). Resilience is considered in a particular clinical practice of handover (Jeffcott et al., 2009b) but the paper does not make specific use of the SWE model. Works found in the literature that make direct use of the SWE model in relation to patient safety are those reviewed below (Amalberti et al., 2006, Cook and Rasmussen, 2005, Miller and Xiao, 2007).

Rasmussen’s (1997) argument is that safety has to be modelled at a higher conceptual level than safety rules and procedures, which leads him to suggest the SWE. There are three constructs to the envelope model: the ‘boundaries’, the ‘operating point’ and the

'gradients'. The envelope has three boundaries: 'Economic Failure', 'Unacceptable Workload' and 'Unacceptable Performance', otherwise described as 'Safety Failure' (see Figure 3.4).

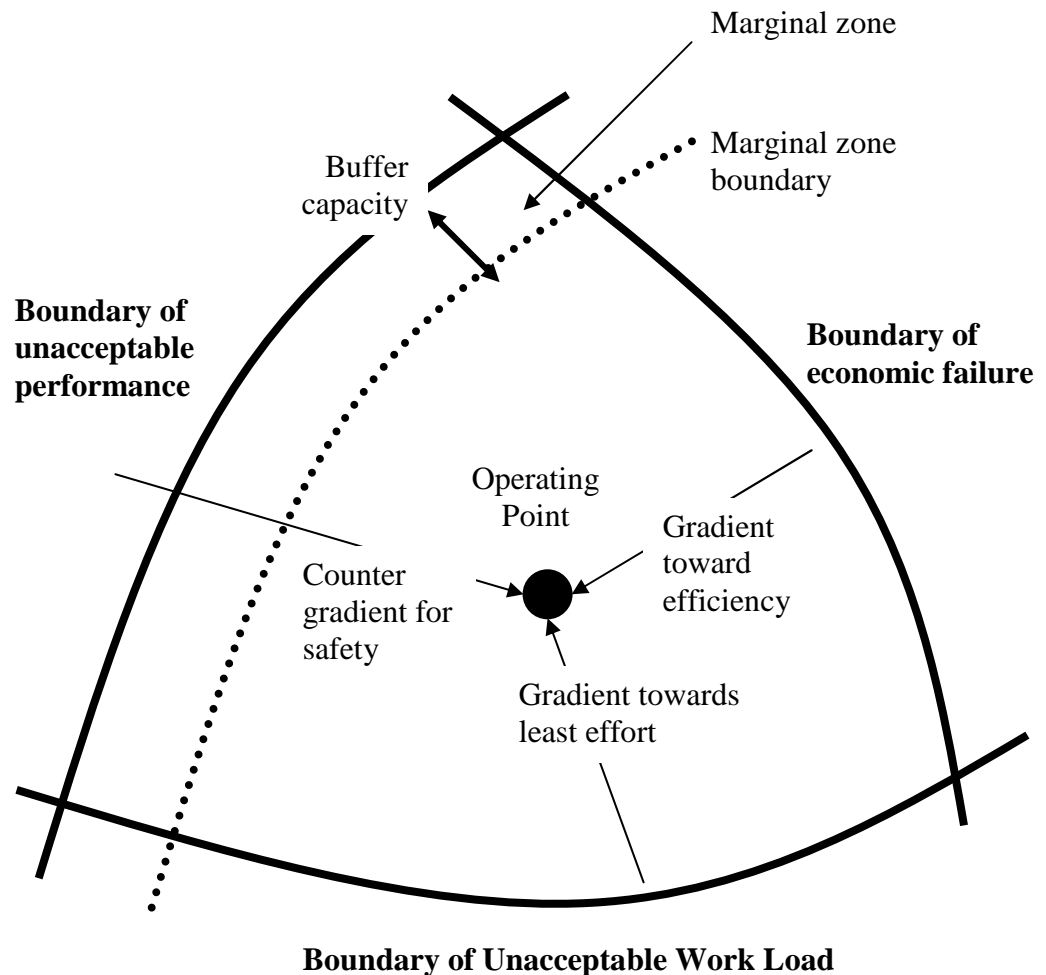


Figure 3.4: Safe working envelope model
(adapted from Rasmussen, 1997)

Part of Rasmussen's (1997) argument is the need to move away from considering risk at the level of tasks. This can be done *'by making the boundaries explicit and known and by giving opportunities to develop coping skills at boundaries.'* (Rasmussen, 1997, p.191 original emphasis) Therefore, he argues that the envelope boundaries have to be made 'visible'.

A 'marginal zone' is depicted in the model as an area close to the safety failure boundary. The zone depicts the system's ability to cope. Therefore, the further away the marginal boundary is from the failure boundary, then there is an increased capacity to

respond to pressures on the system (Miller and Xiao, 2007). From this explanation it is suggested that the marginal zone can be described as the 'buffer capacity' of the system. The resilience of a system is therefore depicted by the buffer capacity. It is argued that the 'location' of the 'marginal zone' boundary is established over time as an organisation sets the limit of what is an acceptable level of risk (Cook and Rasmussen, 2005). Crossing the marginal boundary breaches the social norms of the organisation.

The marginal boundary can 'move'. Inquiries into accidents that breach the unacceptable performance boundary can reposition the marginal boundary inwards (Miller and Xiao, 2007). This is done to provide greater protection (more buffer capacity) to the boundary of unacceptable performance. Equally, over a period of time without accidents, the marginal boundary can 'creep outwards to form a new normal' (Cook and Rasmussen, 2005). The 'boundary' construct dimensions that can be derived from the literature are: 'visibility', 'location', 'buffer capacity' and 'movement'.

The OP depicts the operating conditions of the system in relation to the boundaries (Figure 3.4). The location of the OP in relation to the 'safety failure' boundary does provide an insight into the characteristics of a system. From a conceptual perspective Cook and Rasmussen (2005) identify three types of system:

- (1) stable low risk where the OP moves in small steps and is located well away from the safety failure boundary;
- (2) stable high risk systems (e.g. so called high reliability organisations [HROs]) where the OP is much closer to the safety failure boundary but where safety is achieved by knowing the location of the OP and ensuring any movements are small; and
- (3) unstable systems (low reliability organisations [LROs]) where the location of the OP is less visible and there are large shifts in location both towards and away from the safety failure boundary.

There are two dimensions about the OP construct that can be derived from this literature. They are the 'location' and 'movement' of the OP.

Cook and Rasmussen (2005) point out that there is little if any research 'characterizing the location or movement of the OP of the system or reducing the size of the OP

motions.’ They also suggest that further research into the ‘factors influencing the marginal boundary location’ is needed.

Rasmussen’s (1997) argument is that we need to move beyond the focus on human errors and violations to the rules of work. The SWE model of system behaviour depicts ‘the mechanisms *generating* behaviour in the actual, dynamic work context,’ (Rasmussen, 1997, p.190) (original emphasis). To represent the ‘mechanisms *generating* behaviour’, Rasmussen’s model (1997) uses the boundaries within which the system operates and the influences (known as gradients) on the OP. The influences are conceptualised as gradients exerting ‘pressure’ on the OP. For example, the gradient towards ‘efficiency’ forces the OP away from the ‘economic failure’ boundary. Likewise the gradient of the staff towards ‘least effort’ drives the OP away from the ‘unacceptable workload’ boundary (Figure 3.4).

The model conceptualises the combined pressure from the two influences of ‘efficiency’ and ‘least effort’ forces the OP towards the safety failure (unacceptable performance) boundary and therefore a potential accident (Figure 3.5). The two pressure gradients are met by a ‘counter gradient’ that influences adherence to safety standards. The OP can also be held within the SWE by compensating actions taken by staff (Figure 3.6). The idea of compensating actions that counteract the pressure on the OP towards creating unsafe conditions will be explored further below. The literature points to the gradient construct having a ‘pressure’ dimension.

Another way of thinking about the gradients and boundaries would be to regard them as setting the ‘conditions’ within which the performance of the system takes place. Such conditions are rarely static and therefore, through the movement of the OP, the model depicts the dynamic concurrences that occur (Hollnagel, 2006).

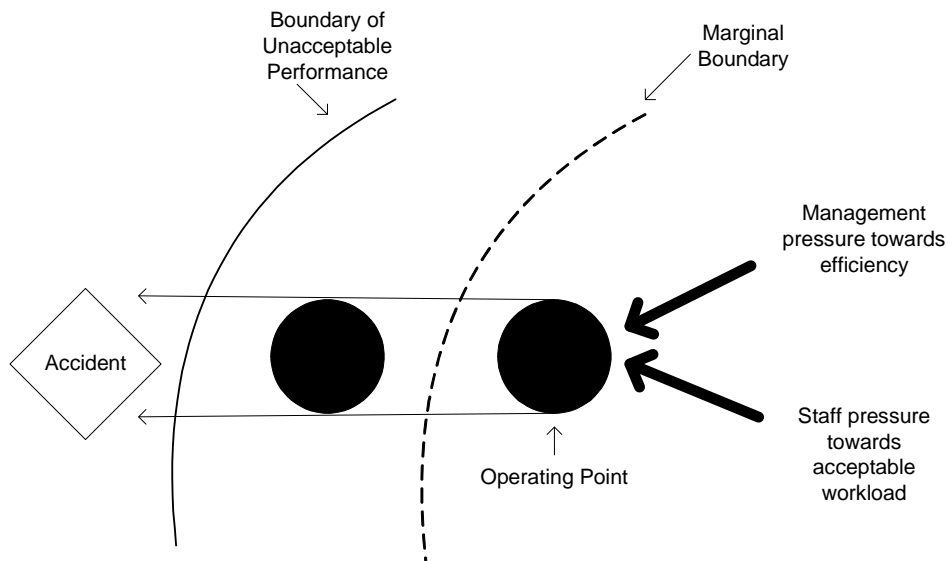


Figure 3.5: Gradients of pressure on the operating point (adapted from Miller and Xiao, 2007)

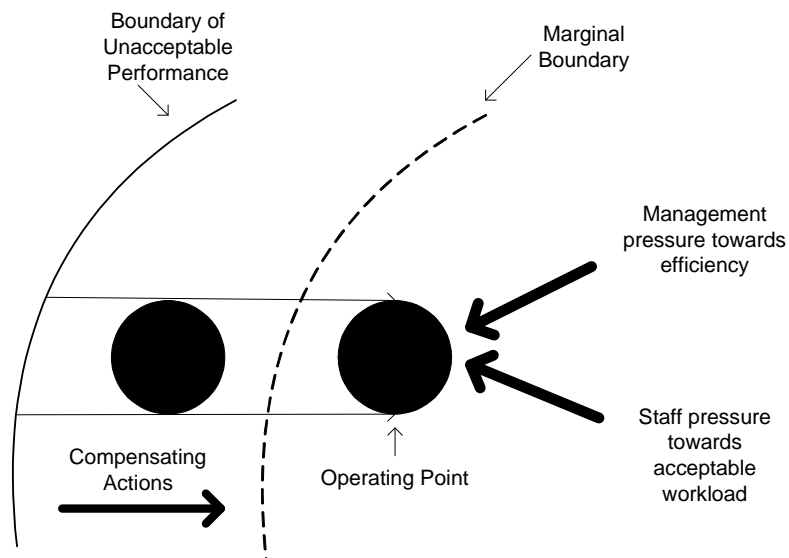


Figure 3.6: Counter gradient to hold the operating point in safe position (adapted from Miller and Xiao, 2007)

The OP can be held within the envelope or moved back out of the marginal zone by the compensating actions or strategies that are in place (Miller and Xiao, 2007). There can be rapid acting compensating actions which return the OP within the SWE. ‘De-compensation’ is what happens when the buffer capacity that has maintained the OP inside the marginal zone, has been exhausted (Woods and Cook, 2006). There are

examples in the literature of where de-compensation occurs quickly and the OP then breaches a failure boundary (Cook and Rasmussen, 2005). This is described by Miller and Xiao (2007) as ‘acute de-compensation’ which is defined as ‘the short-term exhaustion of compensatory mechanisms’. It can cause a rapid change in the dynamics when parts of the system become more tightly coupled and generate new feedback loops and increase vulnerability (Perrow, 1984, Cook and Rasmussen, 2005).

There is another type of de-compensation which is linked to the concept of ‘drift’. In this situation the OP moves closer to the boundary of safety failure slowly over time, as new ‘norms’ are established under financial and production pressures. Miller and Xiao (2007) describe the concept of ‘chronic de-compensation’ where there is a ‘long-term erosion of buffering capabilities at all levels of management’. Where erosion occurs it is depicted as the marginal zone boundary moving closer to the boundary of safety failure leaving little, if any, buffer capacity.

Woods and Cook (2006), suggest that there is a two-stage pattern of de-compensation. The first stage is an automated loop response to the disturbance being experienced. For example, increasing patient admission demand can be regarded in certain situations as a ‘disturbance’ to a hospital system. A number of actions can be taken by staff to compensate for the increase. Such actions include diverting admissions which reduce the continuous pressure on the system (Fatovich et al., 2005, Cook and Rasmussen, 2005). However, the success of that initial response can hide the impact of the disturbance. Such a disturbance may occur repeatedly over a prolonged period leading to chronic de-compensation, to which staff ‘normalise’ (Vaughan, 1996). The second stage occurs when the initial response fails to continue to deal with the disturbance. This stage requires non-automated decision making and new actions to be taken to control the location of the OP relative to the safety failure boundary. Woods and Cook (2006) argue that the success of the first stage can mask how the adaptive capacity is being stretched and the buffers are being exhausted.

Compensating actions are undertaken by staff working in a situation of competing and dynamic pressures. These are conceptualised by the gradients and boundaries in the model (Rasmussen, 1997). Woods (2006) describes part of the dynamic as ‘scale interactions’. His argument is that resilience is influenced either by a ‘downward scale’, where the organisational context create problems for operational staff, or by an ‘upward

scale'. An 'upward scale' is where operational staff adaptations/workarounds and compensating actions influence the more strategic intentions of management. In other words, there are complex dynamics created between the blunt and sharp end of a system that can impact on the OP within the SWE. Therefore, the conditions represented by the gradients provide some insight into the 'pressure' on the OP (downward scale) but the response of staff in the system to those gradients will generate their own 'feedback'. That type of internal feedback (upward scale) is not accounted for in the Rasmussen (1997) model. Such dynamics will influence the resilience of the system. The model therefore needs to be extended to address the impact of the internal feedback dynamics (Hollnagel, 2006) represented by the downward and upward scale interactions. The gradient dimension of 'scale' is therefore derived from the literature.

Woods (2006) argues that we need to identify the different 'classes of dynamic that undermine resilience and result in organisations that act riskier than they realise'. Decision makers are continually confronted with a problem of conflicting goals with many consequential dilemmas and trade-offs to deal with (Hollnagel, 2009a). To choose one side of a dilemma (e.g. production) can create hidden conditions in the system on the other side of the dilemma (e.g. safety) (Reason, 1997). Decision makers do not always examine the consequence, or sacrifice, of their decisions before making them. It is difficult to accurately predict what the result might be, even if time were to be given to such consideration. When a decision is made to increase production the consequences in terms of safety may not be immediately apparent. However, the parts of the system may become more tightly coupled with increased numbers of interactions. The performance of the system, depicted by the OP, may well have moved closer to the boundary of safety failure where a new 'normal' way of working is established.

This moving in an unseen way towards the boundary is a concept that has been termed 'drift to danger' (Rasmussen, 1997) or 'practical drift' (Snook, 2000). Part of this 'drift' is the process of 'normalisation'. Staff can 'normalise' to changes in their working situation. These can include the parts of the system becoming more tightly coupled, having to work harder and faster to meet new production targets. From a safety perspective, 'drift' can be conceptualised as either the OP moving into the marginal zone and therefore closer to an accident, or the marginal unacceptable performance boundary being redefined and moved outwards. When the marginal boundary is moved further out to make the envelope bigger, then it reduces the buffer capacity to deal with

pressure on the OP. Such a movement of the marginal boundary would therefore make the system more vulnerable. The reverse would also apply, where more rigorous safety standards were applied moving the boundary inwards.

Amalberti et al (2006) develops the Rasmussen (1997) model to explain the violation and migration from the designed standard within systems. They argue that the reality of operations create pressures to increase performance leading to individuals cutting corners. The system of operations migrates to an acceptable degree of violation from the designed standards. This situation is termed 'borderline tolerated conditions of use' (BTCUs). It is argued that BTCUs have four features. First, they are regarded as being of benefit rather than a risk. Second, that they improve system performance or help an individual. Third, they are accepted and at times required by managers. Fourth, they are linked to informal rather than formal safety procedures. The social context, it is argued, shifts away from the original design of the system. Over a period of time the violations become custom and practice leading to the 'normalisation of deviance' (Vaughan, 1996).

The migration model is used to derive some key points about safety management in health care (Amalberti et al., 2006). The first is that violations are common, although not many lead to substantial harm. Therefore, they become seen as part of routine work and often become undetectable as violations. Second, violations are poorly understood, difficult to measure and there is little research evidence about their impact. Third, staff can change the social acceptance of the violations and migrate back towards operating to standard procedures. Fourth, violations often occur due to 'trigger conditions' such as over ambitious targets. Staff in such a situation often make a choice between compliance and non-compliance to the standards, as they adapt to the conflicting demands of production and safety requirements.

Cook and Rasmussen (2005) make the point that it is normal for healthcare systems to work at the limit of their capacity, and for Hirschhorn's 'law of stretched systems' to apply:

'every system always operates at its capacity. As soon as there is some improvement, some new technology, we stretch it...' (Cook and Rasmussen, 2005).

Cook and Rasmussen's (2005) argument is that, when systems are under resource and / or performance pressures, the benefits of change are taken in increased productivity or efficiency. The pressure on the OP from the efficiency gradients is described as the system moving back to the 'edge of the performance envelope' (Woods and Cook, 2002).

The key concepts that arise from our review of resilience engineering literature and the SWE in particular are:

- safety is part of a wider socio-technical context which requires a conceptual model to take account of the dynamic interaction of the whole and not just a linear of reductionist view of failure due to faulty component parts of a system;
- a 'SWE' takes account of the wider context and some of the competing dynamics depicted by the three failure boundaries and gradients that influence the OP of the system;
- resilience is conceptualised as the ability of the system to achieve a stable state and remain within the boundaries of the SWE;
- the concepts of 'drift', 'normalisation' and 'trade-offs' can be used in explaining the movement of the OP towards the boundary of safety failure;
- the idea of 'compensating actions' and 'de-compensation' provide an insight into the use of buffer capacity to adapt to a perturbation or continuous stress;
- the location and movement of the OP in relation to the safety failure boundary provides a way to describe the characteristics of the performance of a system;
- there are influences on the OP created not just by the gradients but also the response of staff to the competing dynamics thereby creating upward scale interactions. Those responses create further dynamic interactions that require the Rasmussen (1997) model to be extended.

The SWE has the following constructs:

- 'boundaries' – depicts the constraints within which the system is designed to work;
- 'operating point' – depicts the performance of the system in relation to the boundaries;
- 'gradients' – depict the competing pressures on the OP.

These are theoretical concepts that are not explicitly recognisable for staff working in complex systems. Part of Rasmussen's (1997) argument is that the boundaries need to be made visible to staff. Therefore, a part of this research is to explore the dimensions of these concepts to establish how feasible it is to find ways to depict boundaries, the OP and gradients. The literature points to several dimensions associated with each construct which are summarised in Table 3.1:

Construct	Dimensions	quotes from the literature
Boundaries	Visibility	'making the boundaries explicit and known' (Rasmussen, 1997)
	Movement	marginal boundary 'creeps outwards to form a new normal' (Cook and Rasmussen, 2005) '...IT applications will move the unacceptable performance boundary outwards. The marginal boundary is malleable, however and these gains maybe offset by marginal boundary creep.' (Cook and Rasmussen, 2005)
	Location	'The location of the marginal boundary is determined by sociotechnical processes.' (Cook and Rasmussen, 2005)
	Marginal zone (buffer capacity)	'...the marginal zone as a system's capacity to cope.' (Miller and Xiao, 2007)
Operating Point	Location	'...its distance from the marginal boundary' (Cook and Rasmussen, 2005)
	Movement	'At the moment there is little or no work directed at characterizing the location or movement of the OP of the system' (Cook and Rasmussen, 2005)
Gradients	Scale	downward and upward scale interactions (Woods, 2006)
	Pressure	'continuous pressure' of the safety counter gradient 'compensating the functional pressure of the work environment.' (Rasmussen, 1997)

Table 3.1: Dimensions of the SWE constructs

A summary of how the SWE from resilience theory relates to the concepts derived from accident theory is presented in the next section.

3.4 Synthesis of Resilience and Accident Theory

The definition of ‘resilience’ being used includes the ability of a system to anticipate, adapt and recover from both perturbations and continuous stresses. Resilience and vulnerability are considered to be part of a continuum. In considering resilience in relation to safety of patients within complex socio-technical systems, the three boundary model of a ‘SWE’ is used. Resilience is conceived as the ability of the system in times of disruption, to maintain a stable state and for the OP not to breach the boundaries of the envelope. Such a model provides a means to consider the dynamic nature of complex socio-technical systems as it can take account of external and internal influences. The model moves beyond the simple production / safety trade-off to theorise about the other influences on the performance of the system. Gradients can be linked to the idea of the blunt end influencing the sharp end and potentially creating latent conditions that combine to move the OP through the ‘safety failure’ boundary. However, a limitation of the model is that it does not take account of the potential non-linear responses of staff in the system to the gradient pressures. In other words, the sharp end can exert influence on the blunt end.

The movement of the OP depicts the operating conditions of the system and the relationship to the envelope boundaries. The model includes the concept of ‘drift’ and the idea that people ‘normalise’ to violations, which can be depicted as working close to the ‘safety failure’ boundary. The ideas of compensating actions and the de-compensation of buffer or redundant capacity in relation to the OP, introduces some helpful insights when considering the dynamic nature of safety in relation to other goals in a system. When the OP is held within the envelope, it depicts safety as a ‘dynamic non-event’ (Weick, 1987) where considerable effort goes into avoiding accidents.

The model of a ‘SWE’ provides a system resilience perspective for many of the concepts identified from accident theory. In particular the model can incorporate the concepts of:

- production verses safety – through the boundaries and gradients;
- blunt / sharp end – through the pressures generated by the gradients on front line staff (downward scale interaction);

- latent or hidden conditions – through the ‘conditions’ created by the competing pressures on the OP;
- safety as a dynamic non-event – through the idea of compensating actions holding the OP within the envelope;
- redundancy / buffer capacity – through de-compensation of capacity to hold the OP within the envelope;
- normalisation – through staff accepting the shift in the position of a marginal zone boundary or the OP;
- practical drift – through the gradual movement of the OP or small movements of a boundary;
- trade-offs – through making the boundaries and the location of the OP explicit to decision makers.

The model does not provide a means to explore in depth the concepts of coupling and feedback that result from the conditions generated by the competing pressures on the OP. It is noted that compensating actions and de-compensation can change the dynamics of the system through making the parts more tightly coupled. With tighter coupling the subsequent increasing interactions of the parts generate feedback which in turn changes the dynamic, and potentially, the stability of the OP. Therefore, the concepts of coupling and feedback need to be taken into account by extending the Rasmussen (1997) SWE model. This is done in Chapter 5 using insights from the systems thinking literature.

3.5 Summary

Concepts from the accident theory and systems thinking literature are synthesised with ideas from resilience engineering, and in particular, the SWE model proposed by Rasmussen (1997). It is argued that the SWE can be used as the basis for conceptualising safety within a complex socio-technical system. A resilient system is depicted as one where the OP can remain within the boundaries of the envelope during a disruption or in the face of continuous stress. By using such a system resilience model to guide the empirical data collection of this research, insights are gained about the characteristics of NHS hospitals and how they influence patient safety.

However, there are some weaknesses in the Rasmussen (1997) model that require further developments to be made. It is argued that the model has to be contextualised to assist in understanding the degree of importance given to the boundaries by decision makers. The context also provides insight into the pressure on the OP from the gradients. The internal dynamics associated with the 'structure' (Sterman, 2001) of systems also has to be taken into account.

3.6 The research objective

The literature review identified the need to examine how the 'system' characteristic of healthcare organisations influence patient safety. The result of the review has led to examining the issue from a systemic and resilience perspective. The following gaps in the literature have been identified in ascertaining a system perspective of patient safety:

- a) The exploration of the underlying concepts from systems thinking of what constitutes a 'system' has not been well developed in relation to healthcare organisations.
- b) The extant literature is weak in providing a conceptual model to consider how the dynamics of a socio-technical system can influence patient safety.
- c) The conceptual models used in healthcare situations, such as the Swiss Cheese model (Reason, 1997), are limited as they do not fully take account of the dynamic non-linear feedback that occurs in complex socio-technical systems. The models therefore lack explanatory power.
- d) The conceptual model of a SWE (Rasmussen, 1997) does provide a means to take account of the competing constraints and pressures on a system. However, it does not take account of the 'structure' or 'feedback' that occurs in systems operating inside the envelope.
- e) The SWE model has had limited empirical exploration.

These gaps provide the basis for the objective of the research, which is:

To explore, in NHS hospitals, how a systems approach can inform the development of patient safety theory.

The research takes an empirical theory development approach (Meredith, 1998) to this objective, which includes the development of propositions about improving hospital systems to reduce the risk to patient safety.

Chapter 4 presents the philosophical assumptions and research methodology adopted to operationalise the research. In Chapter 5 the model is developed as a means to guide the data collection and analysis which follows in Chapters 6-8.

Chapter 4 - Philosophical Assumptions and Methodology

4.1 Introduction

This chapter sets out the philosophical assumptions of this thesis, the research design and the methods employed in pursuit of the research objective. The chapter has six sections. The basis for understanding what constitutes knowledge and reality is explored in Section 4.2. The argument is made that taking an objective stance is not possible in complex socio-technical situations. Equally to take a subjective position leaves an entirely relative position. The dualistic subjective or objective approach to knowledge and reality is limited and therefore, the dialectic position of ‘pragmatic critical realism’ (Johnson and Duberley, 2000) is preferred. The consequence of taking a pragmatic-critical realist stance is explored in relation to the ‘structure / agency’ debate, which is at the core of social theory. It is argued that ‘structure’ does influence ‘agency’ in a dynamic rather than deterministic manner.

Section 4.3 looks briefly at the philosophical assumptions found in the literature examined in Chapter 2 and 3. Section 4.4 addresses the methodological options. It is argued that case studies provide the best design to explore the constructs identified in the SWE model identified in Chapter 3. The strengths and weaknesses of using case studies are set out. The selection criteria and details of the cases chosen are given. Section 4.5 sets out the data collection and analysis methods used and how they are deployed to operationalise the research. The ethical issues are identified in Section 4.6 and finally in Section 4.7 the position and influence of the researcher in the research process is acknowledged.

4.2 Philosophical Reflexivity

There are many arguments in the literature about the underlying philosophy of knowledge and reality (Guba and Lincoln, 1994, Kuhn, 1962). One argument emphasises the requirement to set out the research assumptions to provide a basis on which to judge the reliability of the results (Johnson and Duberley, 2000). The context of this research is within what has been described as a ‘complex socio-technical system’ (Rasmussen, 1997). Therefore, in addition to the technical element, the social and systemic aspects of inquiry have to be taken into account. It is argued that social theory

provides the underpinning philosophical position. Various philosophical approaches that seek to explain the nature of social science are explored briefly. The philosophical approach of ‘pragmatic critical realism’ is argued for and applied to the structure / agency debate in social theory.

4.2.1 The nature of social science and management research

Kuhn (1962) argues that a paradigm provides the rules and standards for a particular way of conducting research. A paradigm sets out the assumptions used by researchers about what the world is like (Bryman and Bell, 2007, Guba and Lincoln, 1994, Burrell and Morgan, 1979, Guba and Lincoln, (1994)). Part of Kuhn’s (1962) argument is that research is a ‘strenuous and devoted attempt to force nature into conceptual boxes supplied by professional education.’ Whilst those boxes, known as paradigms, can provide helpful frameworks for the development of science, they can also blind researchers to the possibility of paradigm breaking discoveries (Kuhn, 1962).

Burrell and Morgan (1979) argue that there are four key assumptions that are helpful when studying organisations. These relate to ontology, epistemology, human nature and methodology. They examine these assumptions across the subjective – objective dimension as illustrated in Figure 4.1:

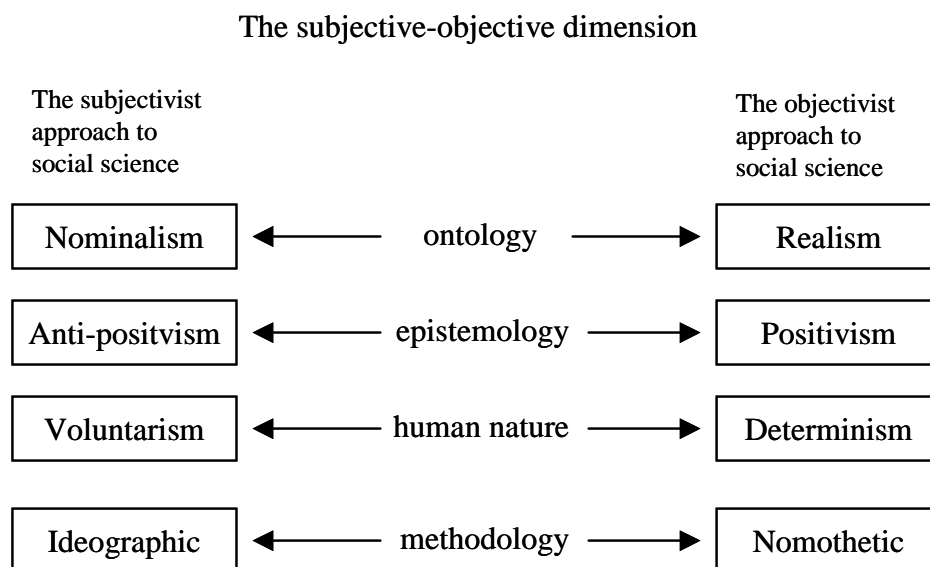


Figure 4.1: Assumptions about the nature of social science – from Burrell and Morgan (1979) p.3.

‘Ontology’ is ‘the general theory about what there is’ (Mautner, 2000). The ‘realist’ position is that there are phenomenon in the world that exist independent from the human mind (Archer et al., 1998, Burrell and Morgan, 1979). The ‘nominalist’ position is that only humans are real and that the social world is a product of the human mind that describes and names the structures and entities in the world (Lane, 1999, Burrell and Morgan, 1979). Others argue that social reality is made up of both items that are human constructs, such as money, and facts. For example the fact that hydrogen atoms have one electron is independent from human minds (Searle, 1995).

Epistemology is ‘the theory of knowledge’ and the ‘scope and limits of human knowledge’ (Mautner, 2000). The ‘positivist’ view of knowledge is that it can only be acquired by objective observation examining reality for generalisable laws (Prasad, 2005). At the other extreme the ‘anti-positivist’ approach is that human interpretation, or personal experience is the basis for knowledge. Knowledge is not revealed through objective observation, rather it is created by the involvement of humans (Prasad, 2005, Lane, 1999).

‘Human nature’ concerns the relationship that humans have with their environment (Burrell and Morgan, 1979). The relationship is also known as the ‘structure / agency’ debate which is at the core of modern social theory (Giddens, 1979). The ‘deterministic’ perspective is that the behaviour of human agents is ‘determined’ by their environment or the ‘social structure’. The ‘voluntarist’ position emphasises the free will of humans and their ability to create and adapt their environments. Some argue that social structures therefore do not exist, seeing them rather as social networks (King, 2004).

The ‘methodological’ strand refers to how knowledge can be obtained. The ‘nomothetic’ approach is to search for generalisable laws with clear concepts that can be measured and tested. The ‘ideographic’ or ‘phenomenological’ perspective is interested in the individual experiences and understanding of the world (Prasad, 2005).

In summary the ‘objective’ end of the continuum of the ‘strands of theory’, lead to a more ‘positivist’ or ‘functionalist’ approach. The ‘subjective’ approach takes a more ‘interpretative’ or phenomenological perspective to social science (Burrell and Morgan, 1979).

Guba and Lincoln (1994), provide a comprehensive review of the different schools of thought. In summary, there are the two approaches of positivist and phenomenology. Within phenomenology they suggest three paradigms: realism, constructivism and critical theory (Guba and Lincoln, 1994). When considering research approaches to theory building, Perry (1998) points to the differences between deductive and inductive. His argument is that positivism takes a deductive approach whilst the phenomenological paradigms use inductive logic to create theory (Perry, 1998). Table 4.1 summarises the position:

Paradigm	Deduction / induction	Objective / subjective	Commensurable / incommensurable
Positivism	Deduction	Objective	Commensurable
Critical theory	Induction	Subjective	Commensurable
Constructivism	Induction	Subjective	Incommensurable
Realism	Induction	Objective	Commensurable

Table 4.1: Categorising the implications for research of the four paradigms

When considering the different assumptions used in management research, Johnson and Duberley (2000) set out a framework of the different schools of thought (Figure 4.2). They highlight the limitations of any such framework particularly drawing attention to the simple binary conditions which are often much more complicated in reality.

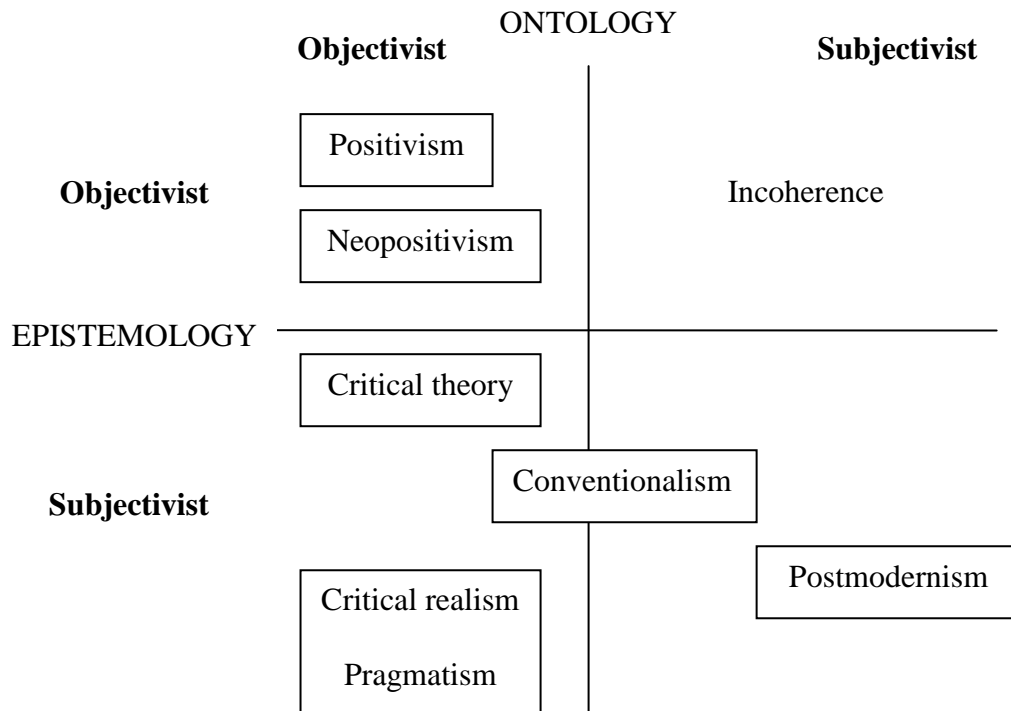


Figure 4.2: Two axis framework depicting schools of management research (from Johnson and Duberley (2000) p.180.)

As noted above, ‘positivism’ is grounded in the belief that it is possible for researchers to be independent and objective observers of facts that constitute reality. It is argued that ‘positivists’ use ‘theory neutral language’. Therefore, reflexivity is confined to the ‘technical’ aspects of the research methods and not to the underlying theoretical assumptions that justify those methods (Johnson and Duberley, 2000, Brannick and Coghlan, 2006). Positivism has a strong background in the natural sciences. Positivism is also criticised by allowing the question ‘what can we know?’ to restrict ideas of what exists (Bhaskar, 1978). Prasad (2005) argues that when studying the social world it is necessary to take account of the ‘capacity for self-reflection and cultural production.’ It is argued that the human ability to interpret differentiates social science from natural sciences (Flyvbjerg, 2001).

‘Postmodernism’ takes a very different approach, although there is little agreement as to what is meant by the term (Mautner, 2000). According to Prasad (2005), ‘postmodernism’ is a radical stance against rationality and meta-narratives that seek to explain reality. It views the world as plural, fragmented and indeterminate. Language is regarded as the source of truth and that images can provide an insight into reality

(Prasad, 2005). What constitutes knowledge is therefore potentially highly individualistic and contextualised (Johnson and Duberley, 2000).

The schools of thought that take the objective ontological and subjective epistemology (bottom left quadrant in Figure 4.2) seek to overcome the weaknesses of both positivism and postmodernism. In simple terms, those schools of thought accept that there is a reality independent of humans but that gaining knowledge about that reality, involves the reflective capacity of researchers. Therefore, a central part of this paradigm is the role of the human agent in relation to their interaction with the independent reality. The critical realist position within this quadrant is explored in further detail below.

It is argued that the realism paradigm is most suitable for this research for reasons that are set out below. In the next section the ‘realist’ position is refined through considering the ‘critical realist’ approach and how that is complimented by ‘pragmatism’.

4.2.2 The Pragmatic-Critical Realist approach

Critical realism takes the view that there is a reality that is independent of human knowledge. As an approach it seeks to understand the ‘enduring structures and generative mechanisms underlying and producing observable phenomena and events’ (Bhaskar, 1989). It is ‘critical’ in that understanding the underlying mechanisms that create observable phenomenon, can lead to changes in the status quo (Bryman and Bell, 2007). The critical realist takes account of the difference between natural and social sciences by recognising that human understandings will vary (Johnson and Duberley, 2000). This recognition requires an interpretive or phenomenological approach.

A detailed study of critical realism is beyond the scope of this thesis. However, there are some key points about critical realism summarised by Johnson and Duberley (2000), which provide the philosophical assumptions underpinning this research.

- a) Critical realists emphasise a metaphysical ontology which states that social and natural reality consists of entities which exist independently of our human knowledge.

- b) The entities may not be observable and different people may apprehend different realities according to varying paradigmatic, metaphorical or discursive conventions deployed through their human agency.
- c) The perceived epistemic role of human agency means that critical realism rejects the possibility of a theory-neutral observational language and a correspondence theory of truth.
- d) Critical realism entails an epistemological defence of causal explanation – causation is not solely expressed through a constant conjunction of events as in positivism. Rather critical realists identify causation by also exploring the mechanisms of cause and effect which underlie regular events.

(adapted from Johnson and Duberley, 2000, p.154)

Critical realism is criticised in a number of ways (Archer et al., 1998, Collier, 1994, Klein, 2004, Lopez and Potter, 2001, Mingers, 2004, Monod, 2004, Johnson and Duberley, 2000). One criticism is that the subjective approach to epistemology makes it hard to justify knowledge claims about processes that are thought to be unobservable, except for examination of their effects in observable events (Johnson and Duberley, 2000). The solution, argues Johnson and Duberley (2000) is for critical realism to adopt aspects from ‘pragmatism’.

‘Pragmatism’ is briefly described as the ‘theory that a proposition is true if holding it to be so is practically successful or advantageous’ (Mautner, 2000). Pragmatists take the view that there is a reality independent of human minds as well as there being a reality within the mind (Creswell, 2007). For example, glass in a window can be observed; ‘glass’ and ‘window’ are linguistic terms that refer to nothing beyond themselves. So it can be argued that the external reality of glass in the window is the product of social discourse. By changing the social construction of language, then it would be possible to deny the existence of glass and the window. Pragmatists argue that there are pragmatic limits to that approach. The pragmatic reality is to try stepping through the window without opening it first, to see if the social construction argument is valid (Johnson and Duberley, 2000).

In summary, Johnson and Duberley (2000) argue that there is a reality that can exist that is beyond our socially constructed discourse:

“...pragmatic-critical realism would argue that while agents socially construct versions of reality through language (interpretative processes from which there is no immunity), the structures of social reality constitute a practical order which acts independently of these constructions so as to constrain or enable our practical actions and interventions.” (p.166)

The practical reality of, for example, glass in a window, constrains the social construction of language to describe that reality. Agreement might be reached to use a different discourse to describe the glass in the window, but the reality remains the same. The pragmatic approach allows critical realism to argue that observing the effect of a directly unobservable phenomenon is a legitimate means of accessing knowledge.

There are two arguments for methodological pluralism that arise from pragmatic-critical realism (Johnson and Duberley, 2000, McLennan, 1995, Creswell, 2007). The first is that different research methods may be required to investigate the structural relations and the human subjective meanings of those relationships. The second argument, based on epistemological subjectivism, suggests that no one method can be regarded as superior. Each method is partial and provides different insights into the reality under examination.

4.2.2.1 Appropriateness of the pragmatic-critical realism

The pragmatic-critical realism is regarded as a subset of the realism paradigm and it is argued to be the most appropriate paradigm for this research. The justification for this position is set out following the categorisation of the different research paradigms (Perry, 1998).

Inductive approach to theory building

A SWE model is developed from the work of Rasmussen (1997), which provides the basis for the data gathering. As noted in the review of the literature, there is little theoretical basis for understanding how and why the dynamic interrelationship in complex socio-technical systems influence patient safety. Therefore, the research seeks to build theory using the SWE model, which describes the phenomenon but is limited in explaining how and why the dynamic interactions occur (Meredith, 1993). To address the research objective it is necessary to examine the interaction between the parts of the

system to find out the characteristics of the hospitals. To conduct the investigation requires the operational definition of key constructs, such as the boundaries of the SWE. How this is done is explored later in Chapter 5. Inferences are drawn from the investigation of social systems with open and fuzzy boundaries rather than hypothesis testing where the concepts can be controlled (Bhaskar, 1979). Therefore, inductive theory building derived from pragmatic-critical realism is regarded as an appropriate approach to study social systems (Riege, 2003, Healy and Perry, 2000).

Objective reality

As noted above, pragmatic-critical realism takes the view that there is a reality that is independent of human knowledge. Unlike positivism, it is accepted that there are aspects of reality that cannot be directly observed and that there will be different perspectives of the same reality. The realist approach to such a situation is to collect data from a number of sources and compare the different perceptions through a process of triangulation (Healy and Perry, 2000, Hammersley, 2008). Therefore, it is argued that researching relationships and their influences in complex social systems from a pragmatic-critical realist position does produce 'observable' data. The data is about the effects of phenomenon that are not directly observable, which produces knowledge (Healy and Perry, 2000). The production of knowledge is the goal of research (Melnyk and Handfield, 1998).

Commensurable

Commensurability requires there to be a common method of measurement (Mautner, 2000). Common measures are required in order to assess the knowledge claims made by the research (Perry, 1998). Perry (1998) argues that the realist paradigm can achieve commensurability by ensuring that the methodology used meets the tests for validity and reliability required for scientific research. These issues are dealt with in Section 4.4.2.1 below.

Practical

The final argument for the appropriateness of the pragmatic-critical realist approach is that it provides a way to engage with the issues faced by practitioners. It is argued that there is a considerable gap between research and practice and that the realism paradigm may assist in providing solutions to problems found in the workplace (Riege, 2003, Bartunek, 2007).

These four arguments justify the basis for taking a pragmatic-critical realist approach to this research. In the next section, the implications of this approach are applied to the debate within social science of the relationship between social structure and agency.

4.2.3 Structure / Agent debate

Before considering the methods employed in this research, it is worth reflecting further on the structure / agency debate in contemporary social theory. This is necessary to understand what is meant by ‘social structure’ at a conceptual level which will have relevance to the application of the SWE model in the case studies.

Giddens (1979) has been at the centre of efforts to reconstitute the idea of social structure with his ‘theory of structuration’. The idea is to move beyond the dualism of structure and action. Giddens attempts to provide a theory that interweaves structure and action in the ongoing activity of social life (Held and Thompson, 1989). Structure is conceptualised as ‘rules and resources’ which provide a framework within which actions take place (Giddens, 1979). There are many critics of Giddens’ approach ranging from his failure to detail what is meant by ‘rules’ (Thompson, 1989), that the rules only provide for ‘cultural’ and not ‘social structuring’ (Porpora, 1998), through to what is perceived as ‘ontological dualism’ (King, 2004).

The social science literature shows no agreement on what is meant by structure or whether such an entity exists beyond social construction (Scott, 2001). An inclusive approach to the analysis of social structure has been advocated by Scott (2001). He argues that social structure is ‘institutional’, ‘relational’ and ‘embodied’. Each of these is explained briefly:

‘Institutions’ form the framework of society through which:

“...practices become culturally standardised and that actions are guided, regulated and channelled. They regulate actions by defining social positions that people can occupy and the behaviour that is associated with them” (Scott, 2001. p.82)

Institutions are cultural phenomena that do not exist in reality. However, structural patterns do exist in ‘the same way as the squareness of a box “exists”’ (Scott, 2001. p.83.)

‘Relational’ depicts the network of social relations between individuals. Specific relationships can be observed, such as a boss to a subordinate. Relationships are dynamic and a process rather than a thing. Scott (2001) suggests that the relationship develops ‘persistent and regular patterns of behaviour’ (p.81).

‘Embodied’ is the term given by Scott (2001) to the work of Giddens (Giddens, 1976, Giddens, 1979, Giddens, 1984) and Bourdieu (Bourdieu, 1977). Embodied structure means that:

“...patterns of institutions and relations are the results of actions on the part of individuals who are endowed with the capacities or competencies that enable them to produce these structures by acting in organised ways. ...Embodied structures are found in the habits and skills inscribed in human bodies and minds. These embodied structures allow them to produce, reproduce and transform their institutional and relational structures.” (Scott, 2001. p.84)

Scott (2001) argues that social structure can be viewed from the three different perspectives of institutional, relational, and embodied as each provides different insights. Such an approach fits with the critical realist position of methodological pluralism and it therefore adopted in this thesis. A summary of the philosophical assumptions used are set out next.

4.2.4 Summary of philosophical assumptions

It is beyond the scope of this thesis to delve into the structure and agency debate in more depth. For the purposes of this research, social structure and agency is regarded as having a dialectic relationship; the agent and the structure are influenced by each other in a dynamic and ongoing manner. By making this assumption the notion that it is the structure that determines the actions of agents is rejected. Equally, the view that structure is purely the product of human agents is not accepted. ‘Structure’ is assumed to have the three aspects of ‘institutional’, ‘relational’ and ‘embodied’ (Scott, 2001).

The philosophical assumptions are based on pragmatic-critical realism. For this research these assumptions are made:

- Objective ontology – there is a reality independent of human knowledge some of which is not directly observable
- Subjective epistemology – there are different interpretations of reality
- Human nature – a tendency towards voluntarism rather than determinism
- Methodology – there are subjective interpretations of reality that require multiple methods to investigate

The pragmatic position is adopted, which is that the key interest is in addressing the research objective rather than on maintaining paradigm purity.

4.3 Philosophical assumptions in the current literature

The philosophical assumptions used in the literature on patient safety, accident theory, resilience and SD is rarely made explicit. There are some clear examples of an ‘interpretive’ approach using case study based qualitative research (Vaughan, 1996, Weick and Sutcliffe, 2001, Snook, 2000). Interestingly, the debate between ‘normal accident theory’ (NAT) and ‘high reliability theory’ (HRT) may have more to do with the underlying philosophical assumptions of the advocates of each theory, than differences in reality. HRT has been derived from an assumption by the researchers of ‘voluntarism’ on the part of the agents who can influence the structure in complex systems. NAT researchers appear to take a more ‘determinist’ position – the complexity of the structure determines the outcome.

Rasmussen (1997) appears to have an objective ontology in setting out a SWE model of a socio-technical system. What is less well developed is the underlying assumption about the role of agents in setting the boundary conditions or maintaining the OP within the envelope. Agents take compensating actions to avoid the OP breaching a boundary. Conceptually these actions appear to be reactive and ‘determined’ by the pressures from the ‘structure’. However, there is dynamic relationship between ‘structure’ and ‘agent’. This means that the ‘agents’ as well as ‘social structure’ are active in determining the location of the boundaries and the pressure on the OP. Agents achieve this through the dynamic feedback relationship with the ‘social structure’.

The assumptions underlying SD are not usually stated. However, some detailed work has shown that if SD is considered to be a ‘method’ rather than a theory, then most applications of SD fit within the ‘functionalist sociology’ paradigm from Burrell and

Morgan's (1979) framework (Lane, 1999). However, a number of papers develop the argument that SD does not fit well within the framework and needs to be situated in social theory (Lane, 2000b, Lane, 2001a, Lane, 2001b). One of Lane's arguments (1999, 2001b), is that SD fits well with the contemporary social theory of the dialectic rather than dualistic relationship between structure and agency. For example, the concept of 'feedback' is core to the dialectic relationship and is a major feature of SD. The assumption is made in this research, following Lane (2001b), that the underlying social theory of SD is the contemporary dialectic relationship between structure and agency, discussed above.

4.4 The research design and methodology

The research design and methods of data collection are dependent on the question to be investigated. From the study of the literature it was recognised that there are a number of gaps in the understanding about the implications for patient safety of the system characteristics such as the dynamic interactions between the parts. At the end of Chapter 3 the following research objective is stated:

To explore, in NHS hospitals, how a systems approach can inform the development of patient safety theory.

To investigate this objective an exploratory and empirical theory building approach (Meredith, 1998, Wacker, 1998) is taken to identify, explain and understand some of the system behaviours of hospitals and their influence on patient safety. There are a number of potential research designs which are briefly outlined. However, it is argued that case study is the most appropriate.

The case studies are used to provide data that is abstracted to populate the SWE. The data provides insights into the context, the planned system design of the hospital, the actual design during perturbations and rich pictures about events, actions and decisions. Inferences and conclusions are drawn about the patterns of system behaviour which provide evidence about the systemic characteristics of the hospitals and how those influence patient safety. Conclusions are also drawn about the development of the SWE

model and the contribution to theory. This high level research process is illustrated in Figure 4.3:

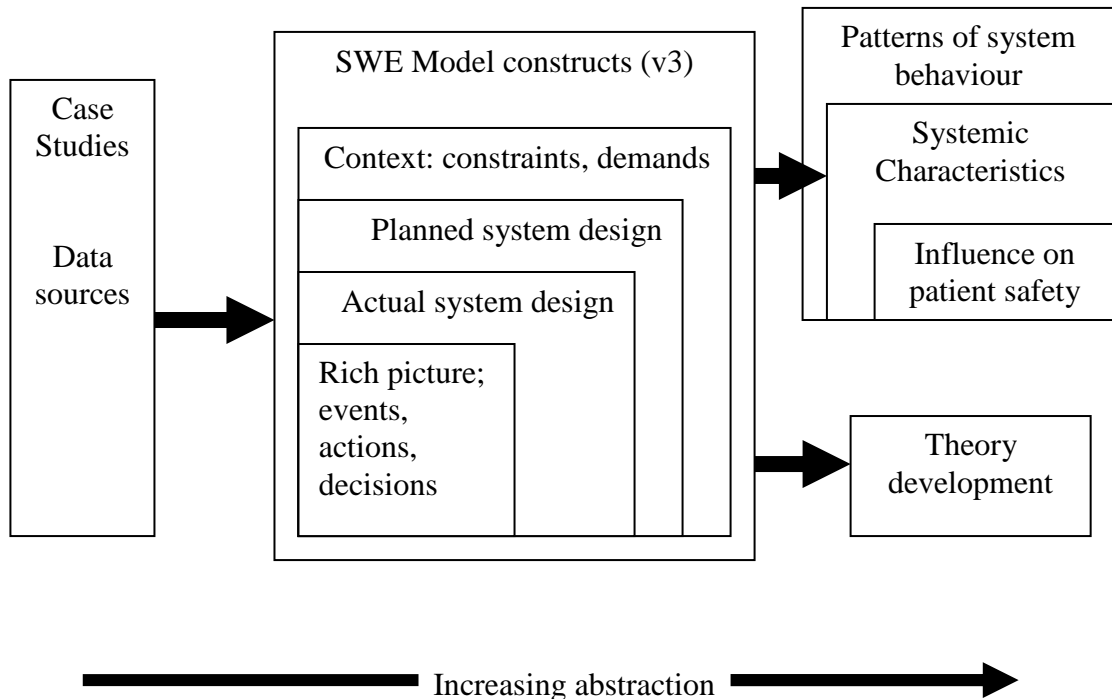


Figure 4.3: High level overview of empirical research process

4.4.1 Research design

The research design sets out the framework for collecting, analysing and discussing the data required in relation to the research objective. There are five main types of design: experimental; cross sectional; longitudinal; comparative; and case study. These designs should not be confused with data collection methods (Bryman and Bell, 2007). These designs are briefly examined to establish the most applicable to the research objective.

Experimental: the basis of this design is to test whether an independent variable makes a difference to a dependent variable. Usually this involves establishing an experimental group and a control group. As Bryman and Bell (2007) point out this type of design is rarely used in organisations due to the difficulty of controlling the many independent variables. This design is not thought suitable for this research.

Cross sectional: this is often referred to as a survey design but other data collection methods can be used. The design entails collecting data at a point in time from a number

of cases (people, organisations) about two or more variables in such a way as to examine and detect patterns from the data (Bryman and Bell, 2007). It requires that the researchers are clear about the variables under investigation and develop suitable data gathering instruments to draw out relational associations. The variables associated with the dynamic interrelationships are not easily determined prior to conducting field work to create a suitable data gathering instrument. This approach excludes the possibility for rich data arising from qualitative methods of interviewing and observation which can provide the means to triangulate the data. Therefore this design is not regarded as the most the most appropriate.

Longitudinal: the object being investigated is studied over a period of time with more than one data collection point. It is often a cross sectional design applied to cases that are repeated. Although there are aspects of this research undertaken over a number of months, the revisiting of data sources is not a key requirement to explore the research question. Additionally, due to the time constraints of this research, this type of design is not considered feasible.

Comparative: the study is conducted using the same data collection methods in two or more contrasting cases. The point is to draw the comparison between the cases investigated. As this research was not seeking to learn from comparisons between hospitals, this design is not considered suitable.

Case study: is an in depth study conducted on one or more cases. Cases can be individuals, organisations, events or locations (Bryman and Bell, 2007, Gerring, 2007). Case studies can be used longitudinally or for comparison between cases. A case study is defined by Yin (2003) as:

“...an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident.” (Yin, 2003) p.63.

The case study design is thought to be the most suitable for this research. The justification for this choice is set out below.

4.4.2 Justification for the use of case study design

A case study is thought to be the most appropriate for this research for the following reasons:

First, the research into the dynamic interactions between the parts of hospital systems requires an understanding of the context within which the complex socio-technical system operates. The case study design is suitable where there is complexity in the subject matter (Stuart et al., 2002). Empirical investigation of cases can provide the richness of understanding that is often not available through other means (Wacker, 1998, Yin, 2003, Rynes, 2007, Voss et al., 2002).

Second, the relationship of the phenomenon being examined to the context of the case study is not known prior to the commencement of data gathering. Yin (2003) argues that case studies are suitable in such circumstance providing that there is a theoretical background to guide the data collection and analysis. The SWE model (Rasmussen, 1997) provides a theoretical basis for this research.

Third, the research takes an inductive empirical approach to theory building (Wacker, 1998). Case study is regarded as a suitable design to answer the what, how and why questions associated with building or developing theory (Meredith, 1998, Yin, 2003, Riege, 2003).

Fourth, there is a desire for the research to have practical application. Using a case study design allows for direct engagement with practitioners and the opportunity for them to comment on the findings and to provide insights that will help them apply the learning arising from the study (Melnik and Handfield, 1998, Rynes, 2007, Voss et al., 2002).

Therefore, case study design is regarded as an appropriate framework through which to gather data to explore the research objective. Before detailing the case studies, data collection and analysis methods used, the issues of the validity and reliability of case studies are reviewed.

4.4.2.1 Judging the quality of case studies

Yin (2003) suggests four tests that can be applied to case studies to judge their quality. These are: construct validity; internal validity; external validity; and reliability. The

objectivist philosophical assumptions behind the need to be explicit on these issues of validity is not always shared by others who take a more subjectivist ontology and epistemology (Stake, 1995, Flyvbjerg, 2006). However, it is helpful to have a means of assessing the strengths and weaknesses of the case study design and to meet the requirement for commensurability (Riege, 2003, Perry, 1998). Yin (2003) suggests a number of tactics about how to conduct a case study that contribute to achieving the necessary validity and reliability, which are summarised in Table 4.2:

Test	Case study tactic	Phase of research in which the tactic occurs
Construct validity	<ul style="list-style-type: none"> • Use multiple sources of evidence • Establish chain of evidence • Have key informants review draft report 	data collection data collection composition
Internal validity	<ul style="list-style-type: none"> • Do pattern matching • Do explanation building • Address rival explanations • Use logic models 	data analysis data analysis data analysis data analysis
External validity	<ul style="list-style-type: none"> • Use theory in single case studies • Use replication logic in multiple case studies 	research design research design
Reliability	<ul style="list-style-type: none"> • Use case study protocol • Develop a case study database 	data collection data collection

Table 4.2: Case study tactics to meet the four design tests
(Yin, 2003, p.34)

The validity of research is concerned about the accuracy of the report of what has been studied (Silverman, 1993). The criticism of case studies is that researchers are too subjective and can lack consistency in establishing what data collection should be undertaken (Yin, 2003). A tactic to improve the construct validity is to use multiple sources of evidence which can then be triangulated (Healy and Perry, 2000). Few other research designs can bring multiple sources of data together on a particular phenomenon (Bryman and Bell, 2007). Therefore, the use of multiple sources can improve both the validity and reliability of the outcome (Hammersley, 2008). Yin (2003) also advises building a 'chain of evidence' and asking key informants from the case study to review the conclusions.

Internal validity is a concern for case studies seeking to make causal explanations (Yin, 2003). Often in qualitative case studies inferences of causation are made. Riege (2003)

argues that case studies, to be internally valid, must not just find patterns but try to show the components which might explain those patterns. Yin (2003) argues that for inferences to have validity, alternative explanations as to why patterns occur, should be considered and the evidence converged.

External validity relates to the generalisable nature of case studies (Riege, 2003, Yin, 2003). Critics often see case studies from the perspective of quantitative research and therefore suggest that the sample size is too small to generate generalisable results. It is often argued that case studies suffer from the weakness of being context specific and therefore transferability of findings to other situations is problematic (Bryman and Bell, 2007). Those arguments misunderstand case study design, which is not a sample from a population (Meredith, 1998, Flyvbjerg, 2006). A small number of cases can raise issues that have wider scale application. For example, Charles Darwin's small scale comparative case study of the differences between the finches in the South American mainland and those on the Galapagos Islands, created a theory on causal mechanism that had wide spread applicability (George and Bennett, 2005)

Flyvbjerg (2006) defends case studies against a number of criticisms. He suggests that properly selected and conducted case studies can be generalisable. He uses the example of the proposition that 'all swans are white', which can be refuted by the observation of one black swan. Flyvbjerg's point is that case studies are excellent at finding 'black swans', as they are often hidden in the detail. Finding 'black swans' can have general significance. He adds that

“...formal generalization is overvalued as a source of scientific development, whereas 'the force of example' is underestimated.” (Flyvbjerg, 2006) p.228.

Yin (2003) suggests one means to improve external validity is having a strong theoretical base for the study or to use the same conceptual framework when doing more than one study. However, as noted above, there are other reasons why case studies can have validity beyond the context of their conduct (Flyvbjerg, 2006).

Reliability seeks to provide evidence of a well conducted study that has minimal errors and biases and could be repeated by another investigator on the same case with the same findings (Yin, 2003). Researching social systems makes any repetition unlikely to produce the same results, but the research procedures should be made explicit (Healy

and Perry, 2000, Riege, 2003). The tactic to help achieve reliability is to create an audit trail setting out the data collection methods used and creating a case study database (Yin, 2003).

The next two sections set out the unit of analysis used and the selection and justification of the case studies.

4.4.2.2 The Unit of analysis

A key principle in the design of case studies is in clearly defining what constitutes the 'case' as a unit of analysis (Gillham, 2000). The research objective is related to hospitals. Therefore, the unit of analysis is the whole hospital as a system. To build a picture of the wider hospital an embedded approach (Yin, 2003), a stratified purposeful method (Miles and Huberman, 1994), is used to examine three within case units of analysis. The reason for choosing an embedded approach is to overcome the weakness of a holistic case study design. Yin (2003) argues that a holistic approach may mean the study is conducted at an abstract level as the researcher may not get into the operational detail. One of the reasons behind selecting three embedded areas is to cover both the formal and informal actions that occur. This assists in understanding the macro context, along with the managers and individual team members (micro) response. The three units are:

Organisational:	Trust Board and hospital wide operational processes
Sub unit:	Division of Medicine
Team:	Ward / Consultant Team / Department members

The departments of paediatrics and obstetrics are excluded from the study as access to those areas was not permitted. Whilst the unit of analysis is the hospital, it is accepted that any hospital works within a wider context that creates influences and responses (see Figure 5.4). The open boundary of socio-technical systems to their context is an important feature that has to be taken into account (Cilliers, 1998). In Chapter 5 the SWE model is developed to include the impact of the context on the operational performance of the hospital and how that might influence patient safety.

4.4.3.1 Selection and justification of the case studies

Miles and Huberman (1994) argue that cases can be selected for a number of reasons. Cases are not a sample from a population. Rather they are selected to provide particular

insights. In this research, having chosen hospitals to be the unit of analysis, purposeful selection (Patton, 2002) was made on the basis of NHS hospitals that provided some contrasting internal characteristics (Miles and Huberman, 1994). The hospitals were also selected within the same Strategic Health Authority (SHA) and Primary Care Trust (PCT) area. This meant that the some of the external context is consistent. A further case study was selected from a different part of England, which provided particular insights into significant patient safety failures. The cases selected are:

Case study 1 (CS 1) was chosen as an example of a high performing large hospital which is a NHS Foundation Trust (FT). In 2007/08 it achieved the highest rating of 'Excellent' for financial management and 'Excellent' for quality of care by the Care Quality Commission (Care Quality Commission, 2009). It has a stable leadership team where the externally validated track record (Healthcare Commission, 2006a) suggests an ability to manage both the external and internal competing pressures. A period of intense pressure on bed capacity during the period October 2008 to February 2009 was studied. This period included episodes of a sickness virus that closed bed capacity and a sustained peak in demand early in January 2009. The points of high demand on bed capacity were examined as a means of magnifying the competing pressures, the staff actions in response to those pressures and the implications for patient safety.

Case study 2 (CS 2) is a small NHS general hospital serving a large rural area. It has a history of financial and operational difficulties with a number of interim chief executives. In 2005/06 it had been ranked the lowest of the four ratings, being 'Weak' for both financial management and quality of care (Care Quality Commission, 2009). In 2007 a new chief executive and chairman were appointed and the financial and operational situation was stabilised. In 2008/09 the Trust was rated second highest out of four possible ratings, being 'Good' for financial management and third highest, being 'Fair' for quality of care (Care Quality Commission, 2009). The Trust seeks to become a FT but has not yet been recommended by the Strategic Health Authority. The period of observation and interviews took place in December 09 – January 10.

Case study 3 (CS 3) is the Mid Staffordshire NHS Foundation Trust. This case was selected as there were known patient safety problems and publically available reports about what happened in the hospital. In 2007/08 and 2008/09 it had been rated 'Good' for financial management and 'Weak' for quality of care (Care Quality Commission,

2009). It is a small NHS general hospital that shortly after becoming a FT was investigated by the Healthcare Commission for an apparently higher than expected level of mortality (Healthcare Commission, 2009a). The results of that investigation led to an Independent Inquiry chaired by Robert Francis QC (Francis QC, 2010a, Francis QC, 2010b). The independent regulator of FTs, 'Monitor', also commissioned an audit report of their processes leading up to the approval of the Trust to become an FT (KPMG, 2009). The management team at Monitor then responded to that audit (Monitor, 2009). The data about CS 3 was derived from the secondary sources listed in Appendix 4.1.

The next section presents the research methods used to gather and analysis the data.

4.5 Research methods

Bryman and Bell (2007) argue that the separation of qualitative and quantitative methods is often over emphasised. The qualitative and quantitative methods often share the subject matter of investigation. As such the clear dividing line between them is questioned when the complicated process of research is taken into account (Bergman, 2008). As noted previously, the context for this research is the complex socio-technical hospital system. Social systems are multi-dimensional (Mason, 2006). It is argued that mixed methods can assist in developing the understanding of complex social worlds. Different methods can provide an insight into both the micro social experience and the macro contextual explanations (Mason, 2006).

Mixing research methods has raised the debate about whether the underlying paradigms can allow such an approach (Tashakkori and Teddlie, 2003, Bryman and Bell, 2007). There are 'purists' who argue that such a mix is not possible and 'pragmatists' who take the underlying assumption that, to gain a wider understanding of the phenomenon, a range of epistemological perspective can be used (Creswell and Plano Clark, 2007). Bryman and Bell (2007) suggest that qualitative and quantitative research do not constitute paradigms given the overlap between them. Yin (2003) argues that by using multiple sources of information it is possible to triangulate the data relating to a phenomenon thereby strengthening the 'construct validity' of the research.

A concurrent approach to gathering data through quantitative and qualitative means is employed to allow different perspectives to be obtained about the hospital being

studied. This helps to increase the validity of the findings (Creswell, 2003). The collection of quantitative data about the changing number of patients in the hospital during peaks in demand informs the wider qualitative study (Creswell and Plano Clark, 2007). The data analysis is an ongoing process of data reduction, display and drawing conclusions (Miles and Huberman, 1994). A summary of the data sources and methods is provided in Table 4.3. Further details relating to each method are set out below (paragraph numbers are referenced in the Tables).

Para Ref	Case	Source of data	Data collection method
4.5.1	All	Documents from: Department of Health; Monitor; Healthcare Commission; Case study hospitals; Primary Care Trust commissioning services from the hospitals; Mid Staffordshire NHS Foundation Trust Inquiries.	Hermeneutical analysis
4.5.1	All	Documents from: Department of Health; Case Studies 1 & 2 Risk Assurance Reports	Content analysis
4.5.2	1 & 2	Hospital staff – purposefully selected	Semi-structured interviews
4.5.3	1 & 2	Hospital staff working – purposefully selected	Non-participant observation
4.5.4	1 & 2	Patient Administration System (PAS); Bed Mgt database	Descriptive statistics
4.5.5	1 & 2	Interviewees	Questionnaire
4.5.6	1 & 2	Interviews; observations;	Casual Loop Diagrams
4.5.7	1 & 2	Interviews; observations;	Stock Flow Diagrams

Table 4.3: Summary of sources of data and data collection methods

The data collected by the various methods is coded and reduced into themes. Table 4.4 details the data collection methods which are used to generate the themes and the paragraph numbers providing further detail below. The details of which data collection method contributed to the themes is also set out in Table 4.4.

Para ref	Data collection method	Theme	Data used
4.5.1	1. Hermeneutical	Context	1,2
4.5.1	2. Content analysis	Planned design	1,4,8
4.5.2	3. Interviews	Actual design	3,4,5,7,8
4.5.3	4. Observations	Rich pictures	3,4
4.5.4	5. Statistics		
4.5.5	6. Questionnaire		
4.5.6	7. CLDs		
4.5.7	8. SFDs		

Table 4.4: Summary of data sources and their application to themes

The data sources used to populate the SWE model constructs and identified characteristics are set out in Table 4.5.

Construct	Characteristics	Data method source used
Boundaries	Visibility	1,2,3,4,5,6
	Movement	3,4,5,7,8
	Location	1,2,3,4,5
	Marginal zone (buffer capacity)	3,4,5,7,8
Operating Point	Location	3,4,5,7,8
	Movement	3,4,5,7,8
Gradients	Scale	1,2,3,4
	Pressure	1,2,3,4,5,6
Structure	Coupling	3,4,5,7,8
Feedback	Type	3,4,5,7,8

Table 4.5: Summary of data sources used to establish the characteristics of the SWE constructs

In the next section further detail is provided about each data collection method set out in Table 4.3 and 4.4 above. The paragraph numbers used below relate to those used in the Tables.

4.5.1 Analysis of Documents

In this research, texts are regarded as part of the reality being studied (Tew, 2001, Flick, 2006). Initially there was a hermeneutical analysis of documents to provide the contextual overview. Hermeneutics is concerned with the interpretation and meaning of texts. There are a number of philosophical positions related to hermeneutics (Prasad, 2005). The assumption made in this research is that to understand the text ('the part') it is necessary to learn about the context ('the whole'). In turn the 'whole' can only be understood from the interaction with and of the 'parts' (Prasad, 2002). This is referred to as the 'hermeneutic circle' in which the readers understanding of the context influences their interpretation of the text and vice versa. Prasad (2005) suggests that such a method can broaden the understanding of an organisation being studied. It is argued that corporate reports or policy documents indicate less about the author than about the concerns of the powerful stakeholders that influence the creation of the text (Prasad and Mir, 2002, Prasad, 2005). Such text may also hold 'hidden' meanings where a particular 'spin' is being placed on the content of the document to hide or emphasise certain realities (Prasad, 2005, Prasad and Mir, 2002). Readers should approach a text both trusting the intention of the document, whilst also being critical (hermeneutics of faith and suspicion) (Prasad, 2005).

'Hermeneutics' in this research is considered as the process of understanding texts from the perspective of the author or stakeholders who create the document (Bryman and Bell, 2007, Prasad, 2002). This requires that attention is focused on the wider context and history within which the documents being reviewed were produced and acknowledging the prior knowledge and expectations of the researcher (Prasad, 2002). Documents from the Department of Health, Monitor, the Healthcare Commission and Inquiry Reports were examined with the prior theoretical basis of the SWE model. The documents were assessed for the emphasis given to themes of finance, targets, staff workload and patient safety, relating to the boundaries of the envelope. The hospital Annual Reports and Trust Board papers were reviewed to provide background information and the importance attached to themes relating to the failure boundaries.

The hermeneutical analysis informed the inferences drawn about the SWE boundary and gradient constructs as outlined in Table 4.5. (The documents collected are detailed in Appendix 4.1)

Quantitative content analysis of certain documents was undertaken (Neuendorf, 2002, Bryman and Bell, 2007). This is done to ‘quantify content in terms of predetermined categories and in a systematic and replicable manner.’ (Bryman and Bell, 2007) The predetermined categories used are developed from the boundaries of the SWE model. Words or phrases associated with each boundary were selected based on previous research and the hermeneutical examination of the documents. The documents were searched for the word or phrase. Where the context of the word or phrase fitted within the predetermined category it was counted. It is suggested that the frequency of occurrence of key words or phrases gives an insight into the value or importance of issues to the organisation producing the document (Bryman and Bell, 2007). The method can also be replicated to study documents published over a time period, as is done in this research.

Allocating words and phrases into categories does require some interpretation by the researcher. The method also has other limitations as it does not examine the wider context within which documents are produced. Therefore, in this research, the quantitative content analysis helps to inform the hermeneutical reading of the documents. Further details are presented in Chapter 5 of how this data source provides evidence for the contextualisation and development of the SWE model.

4.5.2 Semi-structured interviews

During the period of high level demand for inpatient beds, semi-structured interviews of doctors, nurses, managers and directors were conducted and recorded (Kvale, 2008, Rubin and Rubin, 1995, Flick, 2007, Silverman, 1993). The semi-structured interview technique was chosen as it provides for the use of an interview guide, whilst retaining the flexibility to explore deeper into issues raised by interviewees (Rubin and Rubin, 1995). Rubin and Rubin (1995), argue that the design of qualitative interviewing is ‘flexible, iterative and continuous’. For them that means:

‘that each time you repeat the basic process of gathering information, analyzing it, winnowing it, and testing it, you come closer to a clear convincing model of the phenomenon you are studying’ (Rubin and Rubin, 1995, p.46).

A research protocol with interview questions was developed from the SWE (v3) model (as detailed in Chapter 5) and the research objective (Appendix 4.2). An interview guide provides the initial questions and themes, although issues emerge during interviews that required exploration in line with process suggested by Rubin and Rubin (1995).

Interviewees were purposefully selected as expert informants and stratified from the different professions and embedded areas of each hospital (see Table 4.6, 4.7 and Appendix 4.3) (Patton, 2002). Informed written consent was obtained from each interviewee. Interviews lasted between 30 to 75 minutes. A field notebook was used to record immediate reflections after each interview (Silverman, 2005).

Level	Manager	Nurse	Doctor	Totals
Board	4	1*	1**	6
Directorate	7	4	2	13
Ward	0	7	5	12
Totals	11	12	8	31

Table 4.6: Case study 1 number of interviewees by profession and embedded level in the hospital
(* Director of Nursing; **Medical Director)

Level	Manager	Nurse	Doctor	Totals
Board	4	1*	1**	6
Directorate	3	1	1	5
Ward	0	2	2	4
Totals	7	4	4	15

Table 4.7: Case study 2 number of interviewees by profession and embedded level in the hospital

The recorded interviews were professionally transcribed and then checked by the researcher. Copies of the transcripts were sent to interviewees for them to review and amend as necessary. No changes were made by the interviewees. Analysis of the

transcripts was undertaken using NVivo 8. An initial coding tree was devised using the SWE model and the research objective. Further codes were added arising from the analysis of the text (Strauss and Corbin, 1990). The coding tree hierarchies are presented in Chapter 6, which are used for data reduction and analysis of the interview transcripts. Data is presented in Chapters 6 – 8 to provide evidence of how it informs and populates the SWE model.

4.5.3 Non-participant observation

Non-participative observation of staff actions and meetings were undertaken. This method was used to provide an additional source of data, to gain background information for the interviews and to inform the system dynamic diagrams. A non-participant style was chosen to avoid being drawn into the actual work, to allow a broader access to staff and locations, and to avoid complicating the ethical approval process.

The observations included the actions of bed managers, management meetings, senior nurses, doctors rounds, the emergency department (ED), the medical assessment unit (MAU) (also known as the emergency medical unit - EMU), the process of patient admissions, transfers and discharges. Notes were taken at the time of observation in Field Notebooks (FN). Reflections and themes were noted shortly after the field work and links to theoretical concepts, such as ‘practical drift’ of the OP (Snook, 2000), were recorded. Such a note taking method is advocated to increase the reliability of observational data (Silverman, 1993). The observation and interview data informed the creation of the system dynamic diagrams. The observations also provided data for ‘rich descriptions’ of the competing dynamics and the individual stories and reactions (Stake, 1995). Data from observations was used to triangulate with interview and statistical sources to provide evidence in the ‘rich picture’ and ‘actual design’ themes to populate the SWE model.

4.5.4 Descriptive statistics

Hospital administrative data provide the descriptive statistics relating to demand and capacity. (The list of statistical data gathered is set out in Appendix 4.1) This data is used to inform the wider case study by looking at trends over time using run charts and statistical process control charts (Wheeler, 2003). Triangulation of the data is

undertaken between the interviews, observations and the descriptive statistics. For example, one of the key measures used in CS 1 and 2 was the number of medical patients who are accommodated on non-medical wards each day. These patients are known as 'medical outliers'. This statistic is used as an indicator of the ability of the hospital to meet the demand for inpatient beds and is one way to depict the pressure and location of the OP. Many of the dynamics that occur in relation to pressure on the OP happen over very short time periods of minutes and hours. The case study hospitals do not routinely collect the time of day the inpatient admission, transfer and discharge occur. Therefore, the implications of changes in demand at certain times of the day were gathered from observations, interviews and from the work of the information analysts in the hospitals. Historical data from a previous study that had gathered data by the hour was used [CS 1: xls 1.11].

4.5.5 Questionnaire

A simple questionnaire was administered to every interviewee (Appendix 4.4). It was administered at the beginning of the interview and was designed to obtain demographical data and their perception of priority setting at different levels in the hospital in relation to the failure boundaries of the SWE model. This method is used to supplement the interview data and observations and to assist in triangulation particularly in examining the 'boundary' and 'gradient' construct of the model.

4.5.6 Causal Loop Diagrams (CLD)

CLDs are used as part of the data display (Miles and Huberman, 1994). They also provide an insight into the coupling and feedback loops between the parts of the hospital system that were observed and discussed in the interviews. The principles and conventions used in CLDs are set out in Appendix 2.1. In Chapter 8, CLDs are used to illustrate how the parts of the hospital are related. For example, when the admission rate into the MAU rises the admission rate to the wards also rises as the two department are linked in the flow of patients through the hospital. CLDs are used to show the direction of the relationship; in this example the same direction as both rates rise. When the flow of patients does not follow the planned route, then CLDs are able to show the consequences as new feedback loops are created. The diagrams can show where there is a delay in the feedback process. Delays make it harder for decision makers to

comprehend the relationship between a particular action and the resulting and potentially unforeseen consequences (Sterman, 1989, Dekker, 2011).

A key function of CLDs is to demonstrate the type of the feedback loops that occur. As noted in Section 2.4.2, these are either reinforcing or balancing loops. When a system switches between types of feedback loop dominating it creates certain patterns of behaviour, such as oscillation (Sterman, 2000). CLDs provide a means to increase the understanding as to why the hospitals studied oscillate between stable and unstable states.

4.5.7 Stock Flow Diagrams (SFD)

SFDs are used to display the patient flows into and through the hospital and the location and links between the wards and departments (Lane and Husemann, 2008b). In Chapter 8, SFDs are used as a qualitative method (Wolstenholme, 1999). This method is used to model the high level planned design of the hospitals. This is done to illustrate the separation of the flow of medical and surgical patients (Figure 8.3). The impact on the planned design, which occurs due to the change in the flow of patients, is illustrated with a SFD (Figure 8.10). The change in the direction of the flow of patients when stocks (wards) become full creates a coupling of the previously separated flows of patients.

SFDs extend the SWE by conceptualising the dynamic interactions that occur between the parts of a hospital and how those change when there are perturbations due to high levels of demand.

4.6 Ethical issues

There are four main areas of ethical concern: whether there is harm to participants; lack of informed consent; invasion of privacy or whether there is any deception involved (Bryman, 2004).

This study involved participants from the staff of the case study hospitals 1 and 2. No patients were directly involved in the study although staff working with patients were observed. The patient's permission was sought when the researcher was present. To ensure informed consent to participate and that there is no deception or invasion of

privacy, the researcher provided those being interviewed and observed with an information sheet about the research (Appendix 4.5). Written consent was obtained from those interviewed and those being observed for prolonged periods (Appendix 4.6).

In summary the information sheet provided to each participant set out the following:

- purpose and scope of the research
- who is funding and undertaking the research
- that contributions will be anonymised
- that transcripts can be checked by participants
- that interviews or observations can be stopped at any point without reason
- that if patient safety is being found to be seriously compromised putting patients in immediate danger, the researcher has a duty of care to report that to the on-call Director
- how the results will be disseminated

Given the nature of the research, it is feasible that staff actions being inquired about may have shown that patient safety was being compromised. Therefore the following principles were adopted by the researcher to ensure that no harm was being done:

- If during the study the researcher discovers processes, procedures or systems that might be putting patients in immediate danger he will raise it with the staff concerned, explaining his and their duty of care, and break the confidentiality of the research process and make the problem known to the on-call Director. This action would be undertaken on the basis of the need for all parties to learn and that the staff involved would not be blamed.
- If during the study the researcher discovers processes, procedures or systems that might be putting patient's safety at longer term risk he will raise it with the staff concerned and then the Director with responsibility for risk management or the Chief Executive.
- If during the study the researcher discovers areas within the hospital that are exemplars in improving or managing patient safety he will bring that to the attention of the Director with responsibility for risk management having sought the permission of the staff concerned.

- If during the study the researcher encounters staff that become upset or distressed in any way, the interview or observation would cease immediately and the staff member helped to find assistance as appropriate.

The following ethical principles are applied.

- The anonymity and privacy of those who participate in the research is respected and that personal information and research data is kept confidential.
- The researcher is sensitive to the impact of the research on individuals and will seek to minimise any distress caused by participating.
- All published work is anonymised unless otherwise requested and papers will detail the source of funding for the research.
- The storage of electronic data will comply with the requirements of the Data Protection Act.
- Any publications relating to the outcome of the study will acknowledge the financial support from NHS South West and that the researcher had been employed within the NHS as a senior manager including in one of the case study hospitals (1986-90; 92-97).

Prior to the commencement of the research ethical approval was received from the University of Exeter and the local NHS Ethics Committee (Appendix 4.7)

4.7 Reflexivity

Reflexivity is defined as ‘thoughtful, conscious self-awareness’ that examines not only the research methodology but also the interaction of the researcher with the matter being investigated (Finlay, 2002). Given the previous career history of the researcher (as a Chief Executive, Director and manager within the NHS for nearly twenty years, including employment in case study 1 hospital during 1986-1990; 1992-1997) and the participatory aspects in some of the research methods, there is the need for reflexivity (Holliday, 2002). The researcher influences the collection and analysis of the data (Finlay, 2002). It is a subjective process that exhibits properties of co-production (Steier, 1991). Therefore, the researcher must critically reflect on their own assumptions, views and actions related to the area of investigation. Potter (1996) suggests three strategies to achieve reflexivity. The first is to provide a transparent account of the research process and context which is available for audit purposes.

Second, to disclose any assumptions and bias that might influence the data collection, analysis and interpretation. Third, for the researcher to reflect on the methods and take a self critical stance in relation to the data interpretation in the light of prior assumptions and potential bias (Potter, 1996).

It is recognised that the presence of the researcher has an impact on the staff in the organisation being studied. The focus of the study is on patient safety and certain staff emphasised how safe the hospital is. Other staff used the presence of the researcher to highlight particular safety or operational issues. The researcher used triangulation from two or more data sources to balance this bias.

The personal experiences of the researcher in the NHS up to 2007 include the emphasis on achieving targets and financial balance demonstrated by a tough system of performance management. Other experiences include significant patient safety failures in radiotherapy, breast cancer screening and hospital acquired infections. The researcher has also been involved in introducing systems to improve operational performance and patient safety in hospitals.

This thesis does not present detailed reflexive analysis but does draw out in the discussion some of the personal bias that has influenced the research.

4.8 Summary

This chapter defends the suitability of pragmatic-critical realism for this research. The assumption is made that there is a reality independent from human understanding and that knowledge of that reality is derived by humans taking differing perspectives. Such a philosophical approach is suitable to investigate complex socio-technical systems and allows multiple sources of data to be used to gain the breadth of understanding needed. From social science the view is taken that ‘agents’ have a dynamic interactive (dialectic) relationship with social ‘structures’. Therefore, both ‘agents’ and ‘structures’ influence each other in a process of constant feedback.

Case study is regarded as the most suitable research design as it can deal with the complexity and contextual basis of the research objective. The case studies are identified and eight methods of data collection and analysis proposed. The study takes

the approach of embedding the quantitative data within and informing the wider qualitative study. Thematic analysis of the data is undertaken to inform the population of the SWE model. There are limits to the research which are detailed in Chapter 10. The ethical and reflexivity issues are identified and the actions taken specified.

In the next chapter the Rasmussen's (1997) SWE model is developed and contextualised for application within the case study hospitals.

Chapter 5 - Contextualising the Conceptual Model

5.1 Introduction

The position reached at the end of Chapter 2 is the recognition that the extant literature lack explanatory power as to how the ‘system’ influences patient safety. Key features of systems, such as the coupling and interacting feedback between the parts, which creates the behaviour of the whole, are identified from system thinking. In Chapter 3 concepts are derived from the wider accident theory literature that takes account of systemic issues such as interacting parts, coupling and feedback. System resilience theory is suggested as a means to bring together, in a theoretical framework, the concepts from patient safety, systems thinking and accident theory. The SWE (Rasmussen, 1997) is identified as a model that takes account of the complex socio-technical aspects of healthcare with the many competing dynamics and the requirement for safe performance.

At the end of Chapter 3 it is identified that there are some weaknesses with the SWE model. It does not take account of the ‘cross scale interactions’ (Woods, 2006) or the dynamic interactions created by the flow of work through the system. This chapter further develops the SWE to strengthen the model.

The development of the model is done in two ways. The first is through the synthesis of the literature to incorporate the dynamic occurrences within the envelope, presented in Section 5.1. The synthesis involves bringing concepts from SD and social theory, together with the SWE, to create the SWE (version 2) that incorporates concepts from accident theory.

The second development is the contextualisation of the model, set out in Section 5.2. The extension of the model is undertaken by examining the context within which NHS hospitals in England work. The contextualisation is achieved through a hermeneutic and content analysis of NHS policy documents. The result of this phase of the research clearly highlights the importance of performance targets for NHS hospitals. The original model provided by Rasmussen (1997) (SWE v1) is extended to incorporate ‘target failure’ as a boundary within the model. In addition to the ‘boundary’ construct, four other constructs derived from the developed SWE model are identified. The additional

insights into the system characteristics derived through the application of SD are illustrated.

The chapter concludes by arguing that the constructs from the contextualised SWE (v3), and the concepts from the safety literature identified in Chapter 3, provide the conceptual basis to analyse the case study data.

5.1 Developing the conceptual model

It is argued in Chapter 2 that healthcare operates within a dynamic complex socio-technical system (Plsek and Greenhalgh, 2001, Plsek and Wilson, 2001, Braithwaite et al., 2009, Sweeney and Griffiths, 2002). As noted, the extant literature is weak in providing a conceptual model to consider how the characteristics and dynamics of a socio-technical system influence patient safety. It is therefore, necessary to synthesise the literature to develop a suitable conceptual model that takes account of the dynamic and systemic characteristics that influence patient safety.

5.1.1 Synthesising the literature

There are a number of important concepts relating to safety arising from the characteristics of ‘systems’ that are found in the accident theory literature. However, the conceptual models used in healthcare situations, such as the Swiss Cheese model (Reason, 1997), are limited as they do not fully take account of the dynamic non-linear feedback that occurs in complex socio-technical systems (Roelen et al., 2010, Dekker, 2011).

Rasmussen (1997) argues that for us to consider safety at a higher conceptual level, it is necessary to take account of the complex sociotechnical nature of the problem space. His SWE model is used by ‘resilience engineering’ theorists and applied to a limited extent to healthcare (Cook and Rasmussen, 2005, Miller and Xiao, 2007, Amalberti et al., 2006). For the purpose of this thesis, the SWE model provides a means to articulate two underlying themes. First, the model takes into account the complex and competing dynamics that occur in the performance of a system. Second, a resilient system can be conceptualised as one where the operating point is able to remain within the SWE at times of perturbation or continuous stress and is therefore more likely to keep patients

safe. The model therefore provides a ‘system resilience’ perspective to consider safe performance.

The reader is reminded that the boundaries of the envelope depict the constraints within which the system is expected to operate. The performance of the system in relation to the boundaries is depicted by the movement and location of the OP. The dynamics under which the system operates is depicted by the gradients exerting competing pressures on the OP. The SWE model can also incorporate a number of the concepts drawn from patient safety and accident theory.

At this stage there are four observations to make about the model. The first is that Rasmussen (1997), and subsequently Cook and Rasmussen (2005), identify three gradients that impact on the OP. These are the gradients of ‘management pressure towards efficiency’ ‘least effort’ and ‘campaigns for safety’ (see Figure 5.1). It is argued that these three gradients represent a range of variables and are therefore thematic rather than specific. For example, the gradient ‘towards efficiency’ can include a number of specific issues such as costs, income, the use of equipment and speed of working. This means that what might be found in a context specific situation is a multitude of influences which may not always fit neatly into the three thematic gradients.

The second observation is that the envelope needs to sit within a wider context. Such a context has a bearing on the attention paid by the decision making agents to the boundaries and movement of the OP (Williams and Smart, 2009). The context also contributes to the type and strength of the gradients. The context of healthcare is a complex mix of social, political, financial and regulatory factors which result in conflicting and fuzzy policy (Ham, 2009, Klein, 1995). There is the potential for policy decisions to create latent conditions which have unforeseen non-linear interactions, delayed feedback with potential safety consequences. It is argued in Section 5.2 that the context influences the performance and decision making.

The third observation is that the gradients can be a mix of downward and upward influences (‘scale interactions’), considered in Chapter 3, which are part of the complex dynamic (Woods, 2006). Where there is downward influence, the system context influences the OP. Where there is upward influence it originates from the actions of the

staff where, in the face of the downward influences, they make trade-off decisions, find work around solutions or seek local optimisation. Scale interaction between decision making agents and the wider context can be incorporated by taking into account the link between SD and social theory. As noted in Chapter 4, the assumption made in this research is that the underlying social theory of SD is the contemporary dialectic relationship between the social structure and agency (Lane, 2001b).

The fourth and related observation is that the movement of the OP is due to multiple and dynamic factors, not all of which can be understood or predicted (Stacey and Griffin, 2005). Currently, the model helps us to conceive of compensating actions to maintain the OP within the envelope in the face of competing influences. However, the model does not fully include the dynamics of the stocks, flows and feedback loops within the envelope. This means that the Rasmussen (1997) model does not address how the flow of work through the system creates a series of interactions with decision makers who are seeking to balance a number of competing priorities in a complex environment.

To overcome some of the weaknesses observed, the SD approach is used to extend the SWE model to take account of the dynamics of stocks, flows, feedback, delays and the interaction with decision making agents. SD modelling can be used to illustrate the interrelationships between the parts of a system and the potential consequences of a change in those relationships. SD is used to incorporate the concepts of ‘coupling’ and ‘feedback’ between the parts of the system, which in turn influence the behaviour of the whole system. When a system is disrupted or under continuous stress the dynamics generated by the relationships may change with potential safety implications (Cook and Rasmussen, 2005, Perrow, 1984). By depicting the potential change in the relationship of the parts through SD diagrams, further insights can be developed into the movement of the OP in relationship to the boundaries.

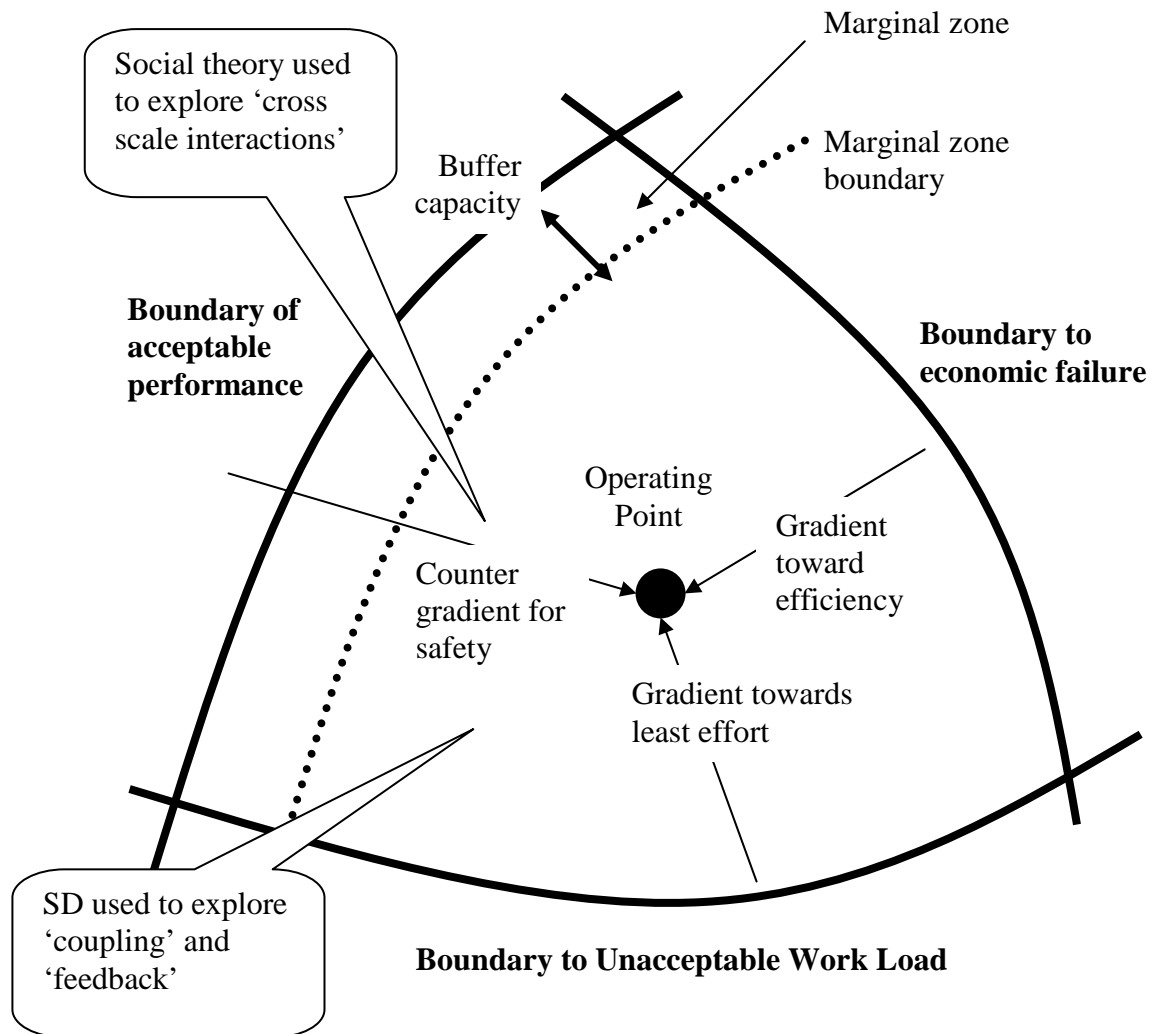


Figure 5.1: SWE (v2) model developed from synthesis of the literature

5.2 Contextualising the Safe Working Envelope model

In Section 5.1, a conceptual system resilience model is developed from a synthesis of the literature. Rasmussen's (1997) interacting three boundary SWE model is extended using concepts from SD to create the SWE (v 2). The version 2 model takes account of the dynamic nature of complex socio-technical systems. The model also depicts the conceptual idea of a resilient system being where the OP of the system remains within the envelope during periods of perturbation or continuous stress. It is argued, that the SWE sits within a wider context. Therefore, to apply the model as a means to conduct research, it has to be contextualised.

The aim of this section is to set the conceptual model within the context of the NHS, and the operational management of hospitals in England. The contextualisation is done by a hermeneutic and content analysis of key documents published annually by the Department of Health in England.

5.2.1 Policy Context

The political and consequent policy context of the NHS influences how the hospitals work (Klein, 1995, Ham, 2009, Walshe and Smith, 2006). In particular, since the publication of the NHS Plan (Department of Health, 2000a) there is an increased emphasis on reducing the wait to access hospital services (Buchanan et al., 2007). There are a number of targets for hospitals to meet to reduce the time patients have to wait (Department of Health, 2005a, Department of Health, 2007, Bevan and Hood, 2006, Buchanan et al., 2007, Buchanan and Storey, 2010).

Each year the Department of Health in England publishes the ‘Operating Framework’, which sets out the priorities and requirements indicating how the NHS will operate in a particular year. To ascertain the balance of priorities given by policy makers, a hermeneutical analysis was undertaken of four ‘Operating Framework’ documents (Department of Health, 2006a, Department of Health, 2007, Department of Health, 2008c, Department of Health, 2006b) (See Appendix 5.1). The hermeneutical analysis shows that ‘targets’ are a major feature in the ‘Operating Framework’ documents alongside financial management. What is not known from that analysis is whether the emphasis on targets is sufficient to justify the explicit inclusion of a dedicated construct within the model and whether this provides explanatory insights into patient safety. To compliment the hermeneutical reading a content analysis was undertaken.

The boundaries in the model are thematic in that they depict a range of influences. A coding protocol was devised based on the words or phrases related to the themes of the SWE boundaries (Neuendorf, 2002). The codes and themes were derived from prior research (Williams and Smart, 2010) and the hermeneutical reading of the ‘Operating Framework’ documents. The documents were searched to identify occurrences of codes relating to each thematic boundary. The semantics of each occurrence of a code within the text was checked to ensure consistency of analysis against each boundary theme. For example, when the word ‘finance’ refers to a job title, such as ‘Director of Finance’, it was excluded. The frequency of each code was calculated and the mean score for each

boundary theme was produced. The summary results of the mean frequency of codes for each boundary theme are presented in Table 5.1 (full results are in Appendix 5.2). From these results, it is noted that the frequency of occurrences of the ‘targets’ code is similar to ‘finance’. The staff workload code is low in occurrence but increasing over time. The number of codes relating to patient safety oscillates, but grows over time. These results suggest that policy makers make greater use of terms relating to targets and finance than the themes from the other boundaries. The substantially increased focus on safety in the 2009/10 document, is the result of the Darzi Review in 2008 (Department of Health, 2008b). This review placed ‘quality’, which includes ‘patient safety’, high on the policy agenda.

Boundary Theme	2006/07	2007/08	2008/09	2009/10
Financial	8.92	13.15	12.62	12.00
Target	6.75	10.38	12.63	12.63
Workload	0.31	2.23	4.69	6.38
Safety	3.4	7.2	5	15.1

Table 5.1: Mean frequency of codes by boundary type from content analysis of Department of Health ‘Operating Frameworks’ for the NHS in England 2006/07 – 2009/2010

The results from this analysis are consistent with previous research, which examines the risk assurance documents (Department of Health, 2003) from four NHS hospitals in England (Williams and Smart, 2010). Whilst this previous research is limited, it is interesting to note the degree of congruence between the documents examined. While the risks that concerned managers in the hospitals are wide ranging, there is a notable emphasis on financial management and achieving targets. Staff workload and safety appeared to receive considerably less attention (Williams and Smart, 2010). These results are in line with accident theory literature; the production orientated boundaries receive more management attention as they are seen as an ‘acute goal’ (Woods, 2006). It is argued that the consequence of focusing on the ‘acute goal’ is that the relationship between the OP and the safety failure boundary is not well understood by decision makers. The importance of targets as a centrally driven means to improve performance is also highlighted in the literature which is considered next.

5.2.2 Extant literature on targets and performance

There is a considerable literature examining performance measurement, management and the use of targets in the NHS and the wider public services in the UK over recent years. A key point that is often made is that what gets measured is what matters (Bevan and Hood, 2006) and those issues that do not have a target can therefore be disadvantaged (Gubb, 2009). The literature suggests that targets and the associated 'star ratings' of services are a public management policy, used extensively by the Labour Government (1997-2010). The often reported aim of developing and using targets is to improve the performance of public services (Bevan, 2006, Hood, 2006, Hood and Dixon, 2010, Barber, 2008, Radnor and McGuire, 2004). Whilst there has been a move towards using a balanced scorecard approach to judge performance (Radnor and Lovell, 2033), personal experience is that certain waiting targets remain at the forefront of performance management in the NHS. However, there is also some evidence presented in the literature of organisations 'gaming' the system and 'hitting the target but missing the point' (Hood, 2006, Radnor, 2008, Mayhew and Smith, 2008). One study seeks to argue that performance dysfunction or gaming does not occur in relation to the 4 hour accident and emergency (A&E) department (Kelman and Friedman, 2009). However, the research takes a reductionist approach and does not take account of the implications to the wider hospital of admitting patients more quickly into inappropriate settings.

Radnor (2008) sets the use of targets in a system perspective arguing that they are designed to provide feedback to the inputs of the transformative process of a public service. In some respects targets achieve the purpose of improving certain performance measures. It has been argued that targets have achieved the required improvement (Bevan, 2009). An example cited by Bevan (2009) is the considerable reduction the number of patients waiting more than six month for inpatient or day case treatment in England (77,000 in 1998 down to 8,000 in 2005). However, as has been noted, targets can create a number of unintended consequences as the interactions generated are not always appreciated. Gubb (2009) argues that targets, such as the 4 hour A&E target, create responses that are detrimental to patient care. He cites the movement of patients to clinical decision units, keeping patients waiting in ambulances, admitting patients unnecessarily and miscoding data. These tactics are used to avoid breaching the target and being shamed by 'politically charged league tables'. Hood and Dixon (2010) suggest that the public, media or academics do not generally have a favourable view of

targets in relation to health or education. Their research found that it is primarily senior civil servants that speak favourably of the use of targets to improve performance.

The targets of interest to this research are those that have a impact on patient safety in NHS hospitals in England. As set out in Section 6.6, there are five specific targets that have implications for the use of bed capacity. Three of these set waiting time targets for certain types of patients; emergency patients in the Accident and Emergency Department (A&E) (4 hours); elective patients to be admitted for treatment from time of initial referral by GP (18 weeks) and cancer patients (one month from diagnosis to treatment). Although these are time targets they translate into production volume targets within the hospital. The production volume targets arise from a combination of the time target and the number of patients in each particular queue (A&E; elective; cancer). The production volume requirement is constrained by the need to admit patients to single sex accommodation which reduces the flexibility of the bed capacity. A further interacting target is the limit (0.8%) of patients who are allowed to be cancelled on the day of their planned admission. Therefore, if a hospital has more than the expected number of emergency admissions it has to use those beds planned for elective patients to accommodate the emergency cases. Yet only a limited number of elective cases can be cancelled so other means of accommodating them are often found. Targets are not just about waiting time and the interactions between them create a set of dynamics within hospitals that needs to be taken into account.

In conclusion, key policy documents from the Department of Health set out a number of key targets, which hospitals must meet. There is also a wide ranging literature acknowledging the use of targets as a public management policy in the UK, and especially in the NHS in England. There is also evidence that targets create unforeseen consequences. It is therefore important, when considering hospitals as complex socio-technical systems with open boundaries to their environment that targets are taken into account.

5.2.3 Developing the model

To reflect the emphasis on targets, the interactive three boundary model (Rasmussen, 1997) is developed by adding an additional boundary of 'target failure' (Williams, 2008). The development builds on the work of Rasmussen (1997), Cook and Rasmussen (2005) and Miller and Xiao (2007). Rasmussen's (1997) boundary of economic failure

is split into the two boundaries of finance and target failure. The four boundaries of the developed SWE (version 3) are: financial failure, target failure, unacceptable workload, and safety failure (Figure 5.2).

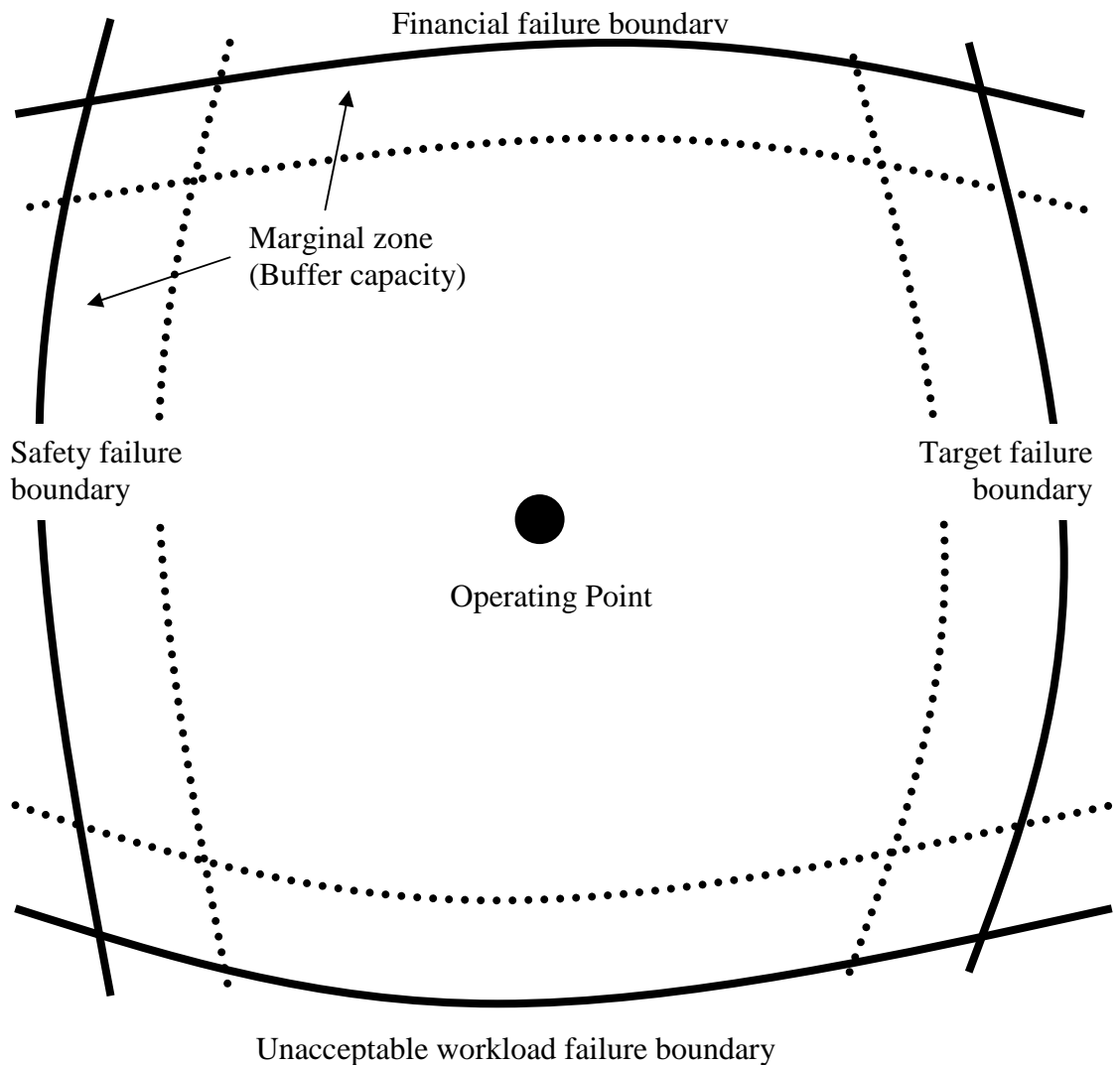


Figure 5.2: The four boundary SWE for NHS hospitals

Rasmussen's (1997) SWE has wide applicability. The developed version (v3) is limited to the specific context of the NHS in England. The developed and contextualised model can be described as a 'mid-range' model (Merton, 1968, Meredith, 1993) in that it is context specific.

The SWE model uses the idea of gradients influencing the position of the OP in relation to the boundaries (Rasmussen, 1997). The gradient depicts the pressure that is exerted

on the OP to keep it away from the boundary. If all the gradient pressures are equal then the OP is held in the mid point of the envelope. The limitation of moving to a four boundary model is that it appears to place the gradients as directly opposing each other (Figure 5.2). However, it is important to note that each gradient can interact with any other. So for example, an increased pressure from the gradient towards an unacceptable workload can move the OP away from the workload failure boundary towards any of the other three boundaries. The direction that the OP moves in will depend on the dynamic of the competing pressures exerted.

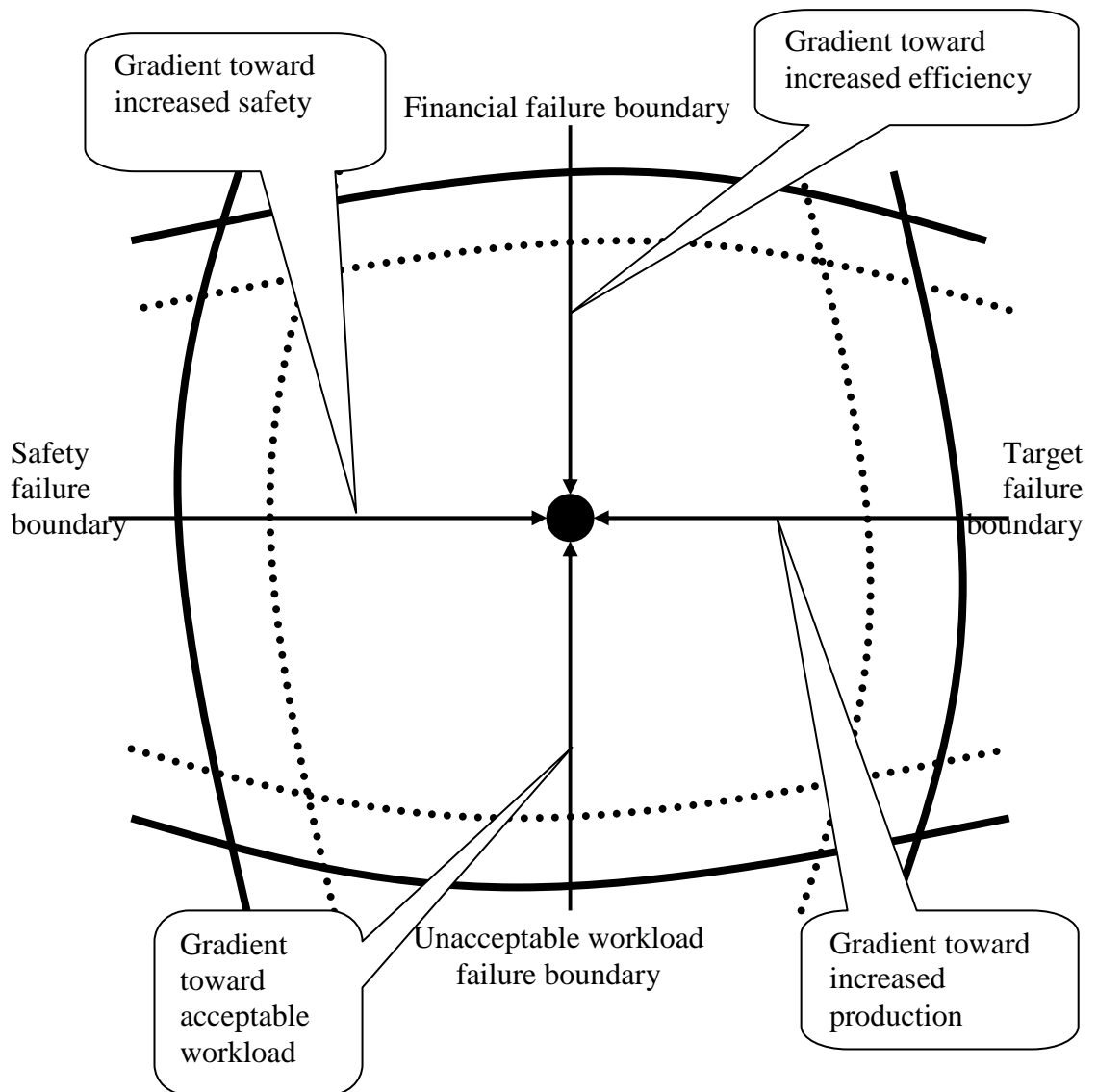


Figure 5.3: Four boundary SWE with gradients

Within the context in which NHS hospitals work there are a number of external influences that contribute to the gradients. The model seeks to show the constantly changing pressures that apply on the OP of a SWE for NHS hospitals. The SWE (v3) is

not exhaustive in depicting the external pressures. Figure 5.4 illustrates that there a number of stakeholders who place sometimes conflicting goals on NHS hospitals (Ham, 2009). For example, there is a strong political and managerial requirement to achieve financial balance or better, whilst at the same time experiencing additional pressure to meet waiting time targets (Department of Health, 2009b). The independent regulator of NHS Foundation Trusts, 'Monitor' sets out clear requirements for meeting both nationally set targets / standards and achieving financial surplus (Monitor, 2008). Similarly, the Care Quality Commission (CQC) sets out the indicators which it takes into account when making assessments of NHS hospitals. These include the national targets / standards (Care Quality Commission, 2010) and the management of financial resources, which is undertaken by the Audit Commission (Audit Commission, 2009). There are nationally negotiated staff contracts that specify the working arrangements for staff, which limits the working hours of key groups, such as junior doctors (Department of Health, 2009a). There is a broader social context where the public has expectations both in terms of access to services alongside assumptions about the quality and safety of the services (Salter, 2004). It is argued that these external influences in a public healthcare system create some of the latent conditions and competing dynamics within which the hospital operates.

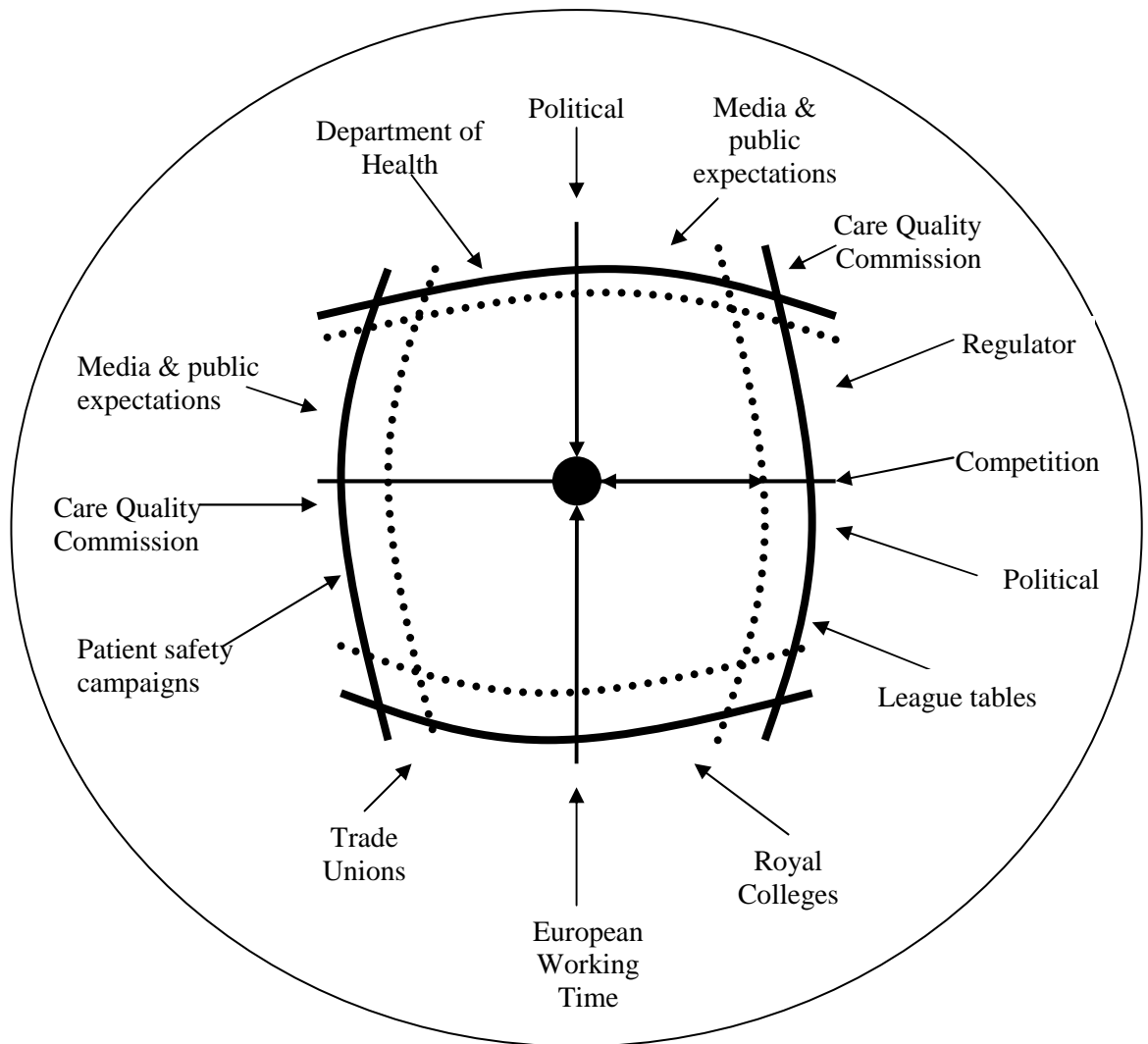


Figure 5.4: SWE set within the wider context of stakeholder influences

It is suggested that these external pressures influence decision making agents within the hospital in both the setting and monitoring of the boundaries. For example, research within the NHS indicates that meeting performance and financial targets is a 'precondition to permit organisations to focus on quality and safety, since the pressures to meet targets compete for senior leadership time.' (Burnett et al., 2010) Equally, there are internal dynamics, which combine with the external influences to create conditions that impact upon the stability and location of the OP in relation to the failure boundaries.

The SWE (v3) model includes four boundaries set within a wider context of influences. In Section 5.1 the original Rasmussen (1997) model is extended by the use of SD and

social theory to take account of the combination of internal and external dynamics within the envelope (Figure 5.5). The construct that is depicted through the use of the SFDs and CLDs from SD is the system ‘structure’. In this context ‘structure’ consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it’ (Sterman, 2001). In a hospital a ‘stock’ depicts the place where patients accumulate, such as a ward or department. The ‘flow’ depicts the direction of movement between stocks.

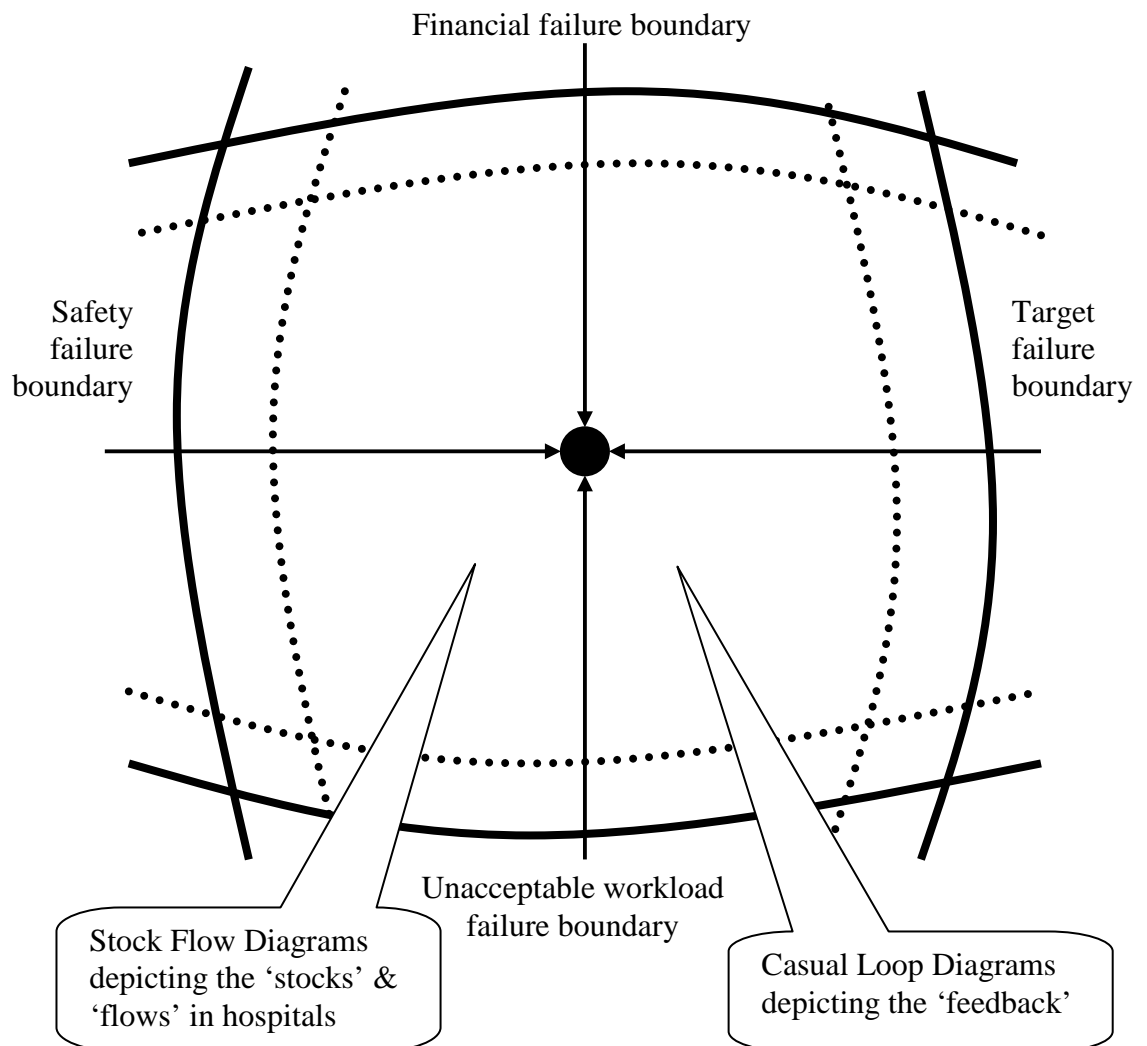


Figure 5.5: Four boundary SWE with SD diagrams to depict system ‘structure’ within the envelope

SD can be applied to the context of hospital operations conducted within the boundaries of the SWE. Set out below are a contextualised SFD and CLD for hospital systems. (In Appendix 2.1 examples are given of the basic conventions used in SFDs and CLDs.)

These are high level diagrams which are developed further in Chapter 8 to display data from the case studies. Figure 5.6 illustrates the basic design of a hospital system for the flow of emergency medical admissions. There are two routes into the hospital. The first is patients attending the Emergency Department (ED) where, once they have been assessed, they are either treated and discharged, or admitted to the Medical Assessment Unit (MAU). The second route in is via a General Practitioner who decides the patient needs direct admission to the MAU. From the MAU the patient is subsequently transferred to a medical ward for ongoing treatment.

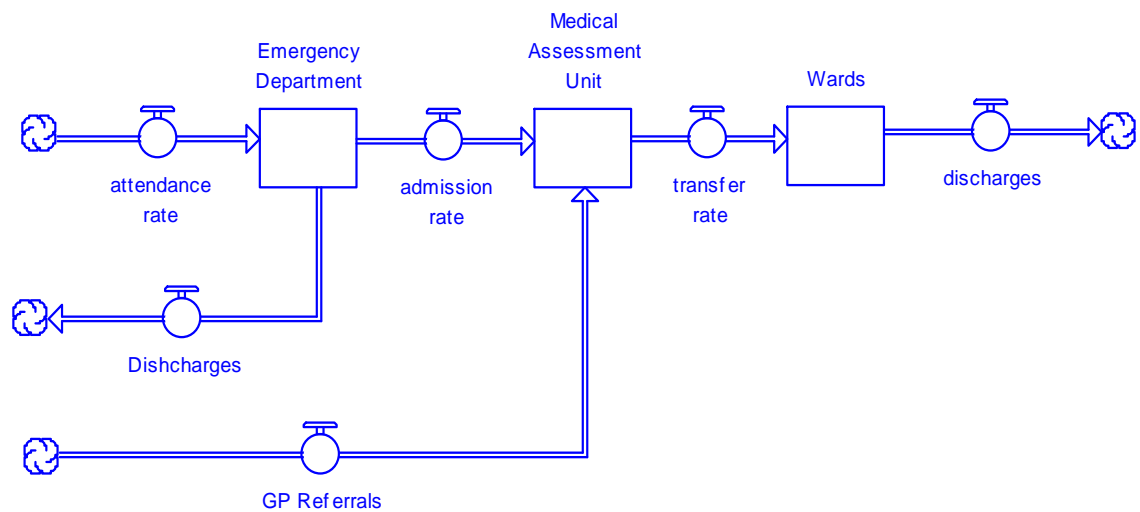


Figure 5.6: Basic stock flow diagram of the emergency medical patient pathway into and through a NHS hospital

SFDs can be used to illustrate changes in flows that can occur when the stocks reach capacity and patients have to be diverted. For example, Figure 5.7 shows an additional flow of diverted patients from MAU to ED. This occurs in hospitals when the MAU has no empty beds to accommodate GP referred patients. Modelling the stocks, flows and feedback loops provide an additional insight into the dynamics within the SWE. The implication of this type of event on the OP is explored in Chapter 7.

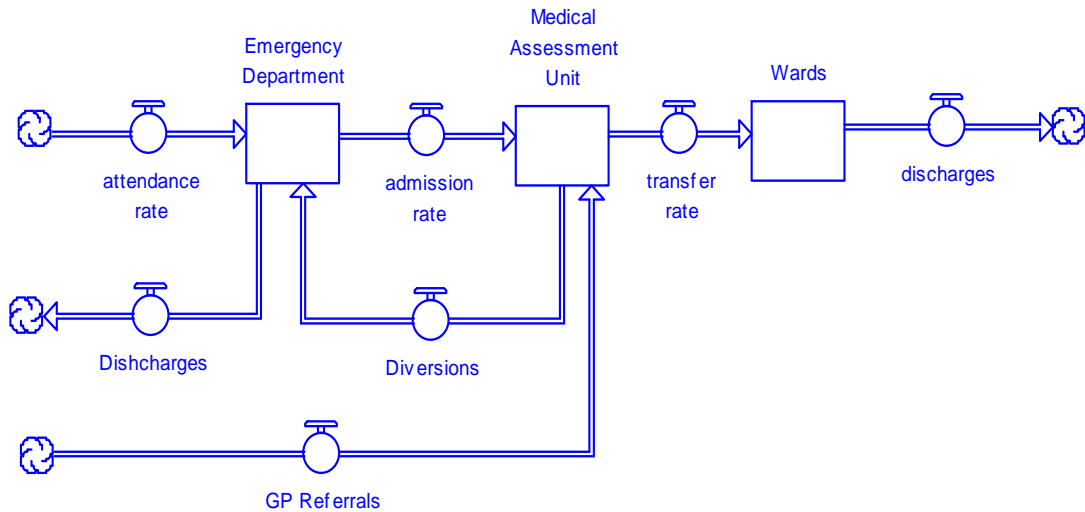


Figure 5.7: Basic stock flow diagram of the emergency medical patient pathway into and through a NHS hospital with diversion flow from MAU to ED

A further insight into the feedback loops that are generated within the ‘structure’ can be illustrated by using CLDs. Figure 5.8 builds on the SFD in Figure 5.7 where there is a flow of diverted patients from MAU to ED. The CLD shows the relationship between the rate of inputs into the hospital system (attendance and GP referral rate) and the occupancy of ED and MAU. The relationship is in the same direction – if the referral rate increases the occupancy increases and vice versa. A reinforcing feedback loop is created between the ED and MAU when patients are diverted. This means that when there are diversions from MAU to ED the occupancy in ED will continue to rise until some balancing feedback loop is initiated. Such a balancing effect can occur by increasing the rate of discharges. Speeding up discharges may have safety implications as research indicates a high rate of adverse events due to poor communication and handover of patients leaving hospitals (Forester et al., 2003, Forester et al., 2004, Kripalani et al., 2007).

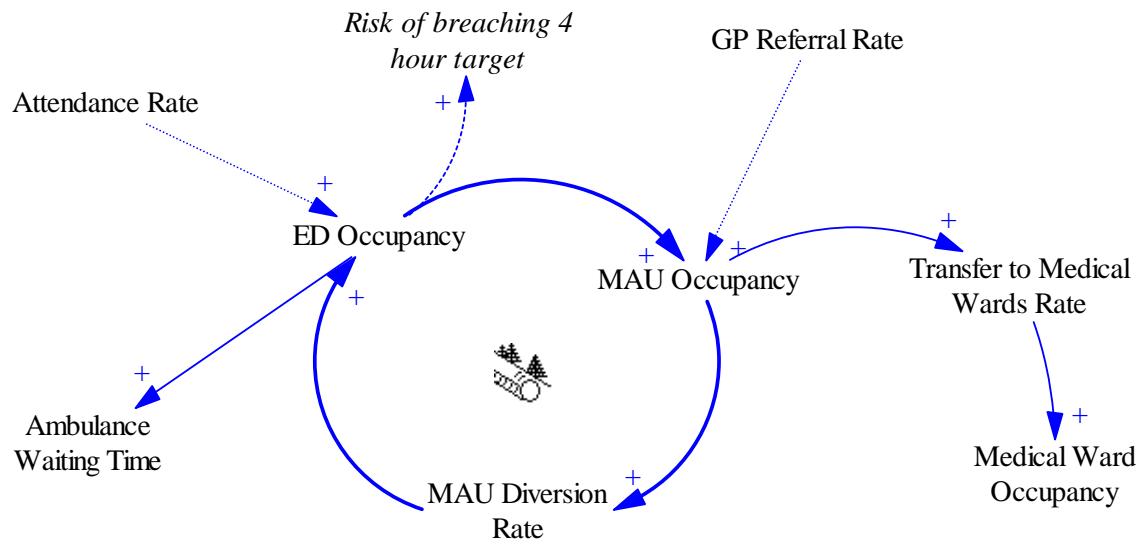


Figure 5.8: Causal Loop Diagram showing the relationship of the emergency medical patients rate of arrival to a hospital system

A CLD can depict the consequences for other parts of the system, including the relationship to policy achievement. In this example, the link is to meeting the 4 hour ED waiting time target. When the ED occupancy increases, then the risk of breaching the target increases and vice versa.

CLDs are used in this research to illustrate the changes that occur in the type of feedback loops that are found within a hospital when it faces continuous stress or perturbation. Such diagrams also help to show how decision makers seek to balance the system, whilst experiencing competing pressures from the gradients, by taking compensating actions to avoid the OP breaching a failure boundary.

5.3 Applying the model

In applying the SWE (v3) to the empirical research it is necessary to clarify the constructs being used. It is argued that the SWE (v3) has the following five constructs:

- ‘boundaries’ – depicts the constraints within which the system is designed to work
- ‘operating point’ – depicts the performance of the system in relation to the boundaries
- ‘gradients’ – depict the competing pressures on the OP
- ‘structure’ of the system and the relationships between the parts

- dynamic ‘feedback’ between the parts that contribute to the stability of the whole system

The term ‘constructs’ is defined in this thesis as ‘an abstract form of concept which cannot be observed directly or indirectly but can be inferred by observable events’ (Meredith, 1993). This approach is in line with the pragmatic critical realist position adopted in this research. The five constructs derived from the four boundary SWE (v3) are explored empirically through the case studies. The investigation gathers data about the construct dimensions which were identified from the literature in Chapter 3. The aim of using this SWE (v3) model is to gain insights into the characteristics of the hospitals studied and how they influence patient safety. The five constructs are grouped into three sets which interact which create emergent system behaviour (Figure 5.9). These are examined in detail in Chapters 6 – 8.

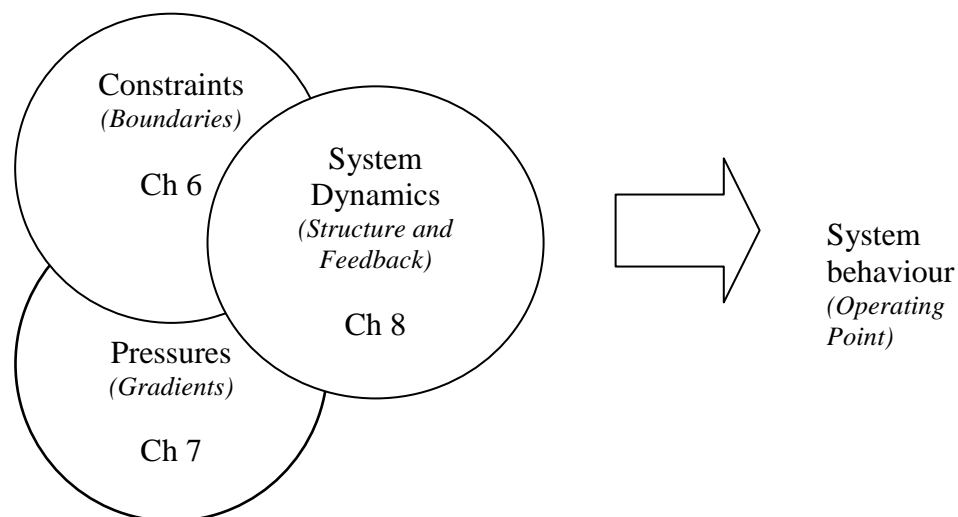


Figure 5.9: Combination and interaction of construct sets depicted by the SWE model

The SWE model is used in healthcare (Cook and Rasmussen, 2005) to consider the consequences for a hospital ‘going solid’ due to a bed crisis. The change in dynamics that occurs when a system becomes tightly coupled, for example, due to lack of bed capacity, is explored in this research through the application of the model. Cook and Rasmussen (2005) have not undertaken an empirical inquiry using the SWE in hospitals. The operations management literature on NHS hospitals and policy documents suggests a number of actions to both reduce and cope with peaks in demand in order to avoid a bed crisis (Klassen and Rohleder, 2001, Armitage and Raza, 2002,

Ham et al., 2003, Proudlove et al., 2003, Proudlove et al., 2007, Department of Health, 2005b, Department of Health, 2010c). However, the link between the operations management of patient demand and patient safety is not usually made.

It is argued that the SWE (v3) can be used to explore a number of concepts derived from the patient safety and accident theory literature that apply to systems. Using the safety theory concepts assists in explaining the influence that the system characteristics have on patient safety. The eight accident theory concepts, identified from the literature, are explicitly linked to the model in the following way:

- the tension between production verses safety – through the boundaries and gradients;
- blunt / sharp end – through the pressures generated by the gradients on sharp end staff and system performance depicted by the movement of the OP;
- latent or hidden conditions – through the ‘conditions’ created by the competing pressures on the OP;
- safety as a dynamic non-event – through the idea of compensating actions holding the OP within the envelope;
- redundancy / buffer capacity – through de-compensation of capacity to hold the OP within the envelope;
- normalisation – through staff accepting the shift in the position of a boundary or the OP;
- practical drift – through the gradual movement of the OP or small movements of a marginal zone boundary;
- trade-offs – through making the boundaries and the location of the OP explicit to decision makers.

The boundary construct is explored in Chapter 6 through analysis of data from the three case studies. The gradients and OP are examined in Chapter 7 using thick description of three events in case studies 1 and 2. The design and implication of the ‘structure’ and ‘feedback’ is presented in Chapter 8. The analysis seeks to conceptualise the findings in terms of both the model constructs and the safety concepts. This is done by indentifying the different SWE construct dimensions and how they incorporate the accident theory concepts.

5.4 Summary

In this chapter the interactive three boundary SWE model (Rasmussen, 1997) is first extended using concepts from SD. Secondly the model is contextualised and developed with the additional boundary of 'target failure'. This is included as a result of the hermeneutical and content analysis of the NHS 'Operating Framework' documents. Both experience and the analysis indicate a high level of attention to finance and targets and substantially less to staff workload and patient safety. The SWE (v3) depicts the context for NHS hospitals in terms of the failure boundaries and gradients that create dynamic influences on the OP. The safety concepts derived from the accident theory literature can be taken into account when using the SWE model.

The five constructs of the SWE (v3) provide a conceptual basis on which to examine the system characteristics of hospital systems. The following three chapters combine the constructs of the SWE with the concepts derived from the safety literature to analyse empirical case data collected during periods of high demand for inpatient beds.

Chapter 6 – Investigating the Boundaries of a Safe Working Envelope

6.1 Introduction

The aim of this chapter is to explore the dimensions of the ‘boundary’ construct of the SWE (v3) model developed in Chapter 5. Each boundary within the model is informed by accounts of the competing constraints experienced by staff. The data from the interviews and observations (CS 1 and 2) and inquiry reports (CS 3) is triangulated with the analysis of documentation, such as reports or policy papers from the Department of Health. The hospitals were studied during times of high demand for their inpatient services where staff managed complex interactions to keep the hospital functioning. Such situations provided data about how the prioritisation of competing demands was managed. While the boundaries are not a directly observable phenomenon in themselves, observable data relating to the articulation, measurement, and prioritised actions of staff relative to each boundary theme was studied.

The ‘boundary’ construct is part of the wider SWE model. The boundaries seek to depict the constraints within which the system operates. In Chapter 7 the pressures, depicted by the ‘gradients’ on the OP are considered. Chapter 8 explores the dynamics that occur inside the SWE generated by the interaction of demand and capacity with decision makers. The data presented in Chapter 7 and 8 provides an insight into the position and movement of the OP. The three chapters set out the data from the case studies in such a way as to suggest that it is the combination and interaction of the three construct sets (constraints, pressures, the dynamics of demand and capacity with decision makers) that provide an insight into the behaviour of the system in relation to patient safety (Figure 6.1). This Chapter focuses on the ‘constraints’ depicted by the boundaries of the SWE and how they influence the behaviour of the system. It is recognised that the ‘constraints’ interact with the other construct sets which are explored in later chapters.

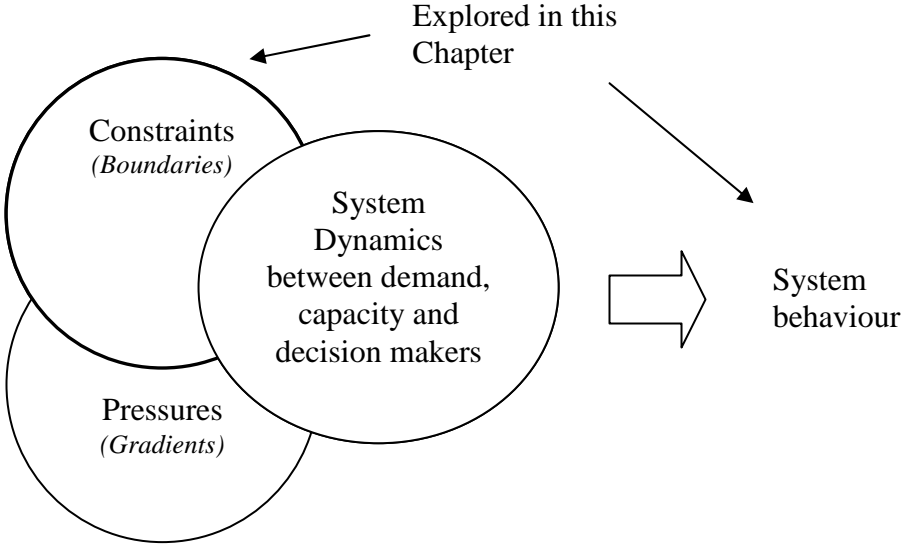


Figure 6.1: Combination and interaction of construct sets depicted by the SWE model

The process of data reduction, display and conclusion is undertaken in an ongoing rather than linear process (Miles and Huberman, 1994). The data is reduced and displayed in relation to each boundary. The coding tree hierarchies are presented below for the SWE constructs and themes of ‘actual design’ and ‘rich pictures’ which are used for data reduction and analysis of the interview transcripts (Figures 6.2 – 6.4).

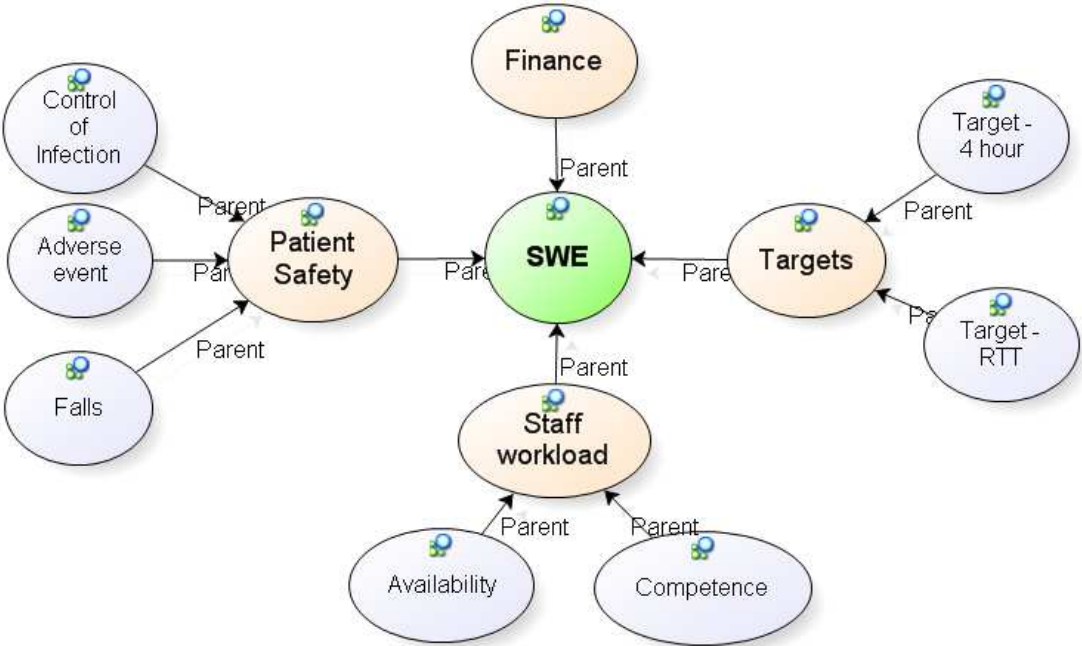


Figure 6.2: Coding hierarchy for ‘Safe Working Envelope’

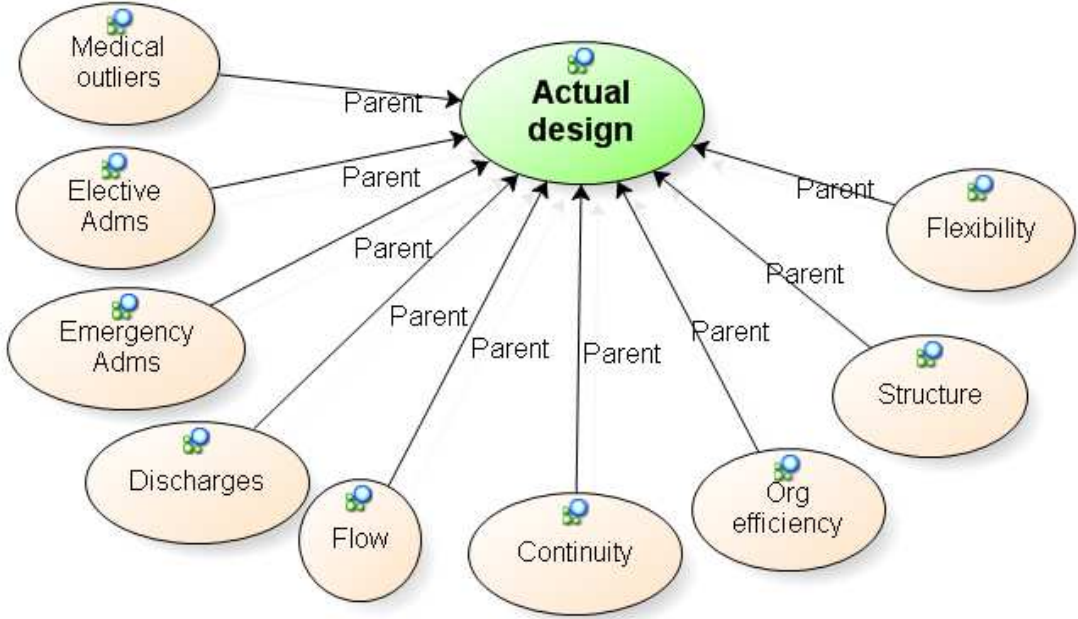


Figure 6.3: Coding hierarchy for 'Actual Design' theme

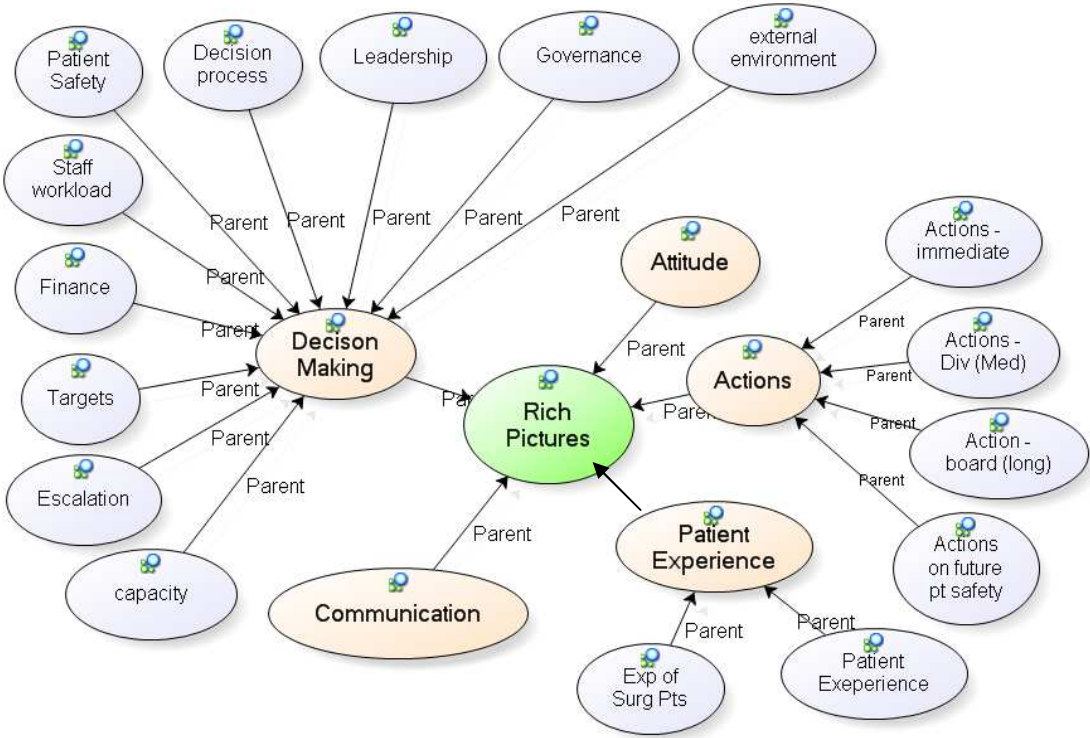


Figure 6.4: Coding hierarchy for 'Rich Pictures'

The themes coded to ‘rich pictures’ and ‘actual design’ contribute to populating the SWE (v3) model. This is done in conjunction with data derived from the document analysis which examines the ‘planned system design’ and ‘context’ (see Figure 6.5).

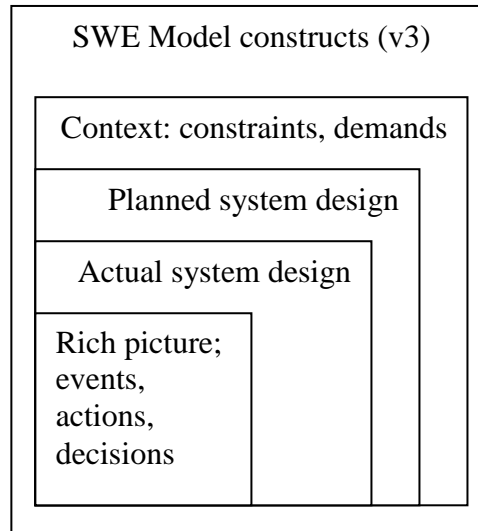


Figure 6.5: Relationship of themes to SWE (v3) model

Sources of data are presented inside square brackets []. The sources refer to observations notes in the Field Notebooks [FN showing the book and page number] (see Appendix 6.1), coded data (see Appendix 6.2); document or statistical data number [see Appendix 4.1), or interview transcript [case study interviewee number] (see Appendix 4.3).

Dimensions of the boundary construct of the SWE are identified from the literature in Chapter 3. For ease of reference, these are presented again in Table 6.1.

Construct	Dimensions	quote from the literature
Boundaries	Visibility	‘making the boundaries explicit and known’ (Rasmussen, 1997)
	Movement	marginal boundary ‘creeps outwards to form a new normal’ (Cook and Rasmussen, 2005) ‘...IT applications will move the unacceptable performance boundary outwards. The marginal boundary is malleable, however and these gains maybe offset by marginal boundary creep.’ (Cook and Rasmussen, 2005)

	Location	‘The location of the marginal boundary is determined by sociotechnical processes.’ (Cook and Rasmussen, 2005)
	Marginal zone (buffer capacity)	‘...the marginal zone as a system’s capacity to cope.’ (Miller and Xiao, 2007)

Table 6.1: Dimensions of the SWE ‘boundaries’ constructs derived from literature

The data is used to investigate the dimensions identified (‘visibility’, ‘movement’, ‘location’, and ‘buffer capacity’). The interrelationship between the boundaries, gradients and OP is also of importance, which is discussed in Chapter 7. Before presenting the data on each failure boundary Section 6.2 sets out the results from the questionnaire administered to the interviewees to gain their views about the priorities given to the competing goals.

6.2 Staff views on the priorities of the boundaries

Interviewees in Case Study 1 and 2 were asked, using a simple questionnaire (Appendix 4.4), to give their view on how the trust board, divisional management teams, and clinical staff teams on wards, ranked the priorities of patient safety, finance, targets and staff workload. The interviewees were asked to rank the different priorities in order of importance for decision making in the hospital. The results, based on the percentage ranked the highest priority for each level are shown in Tables 6.2 and 6.3.

	Achieving Targets %	Adequate staffing	Patient safety	Achieving financial results
Trust Board	15%	0%	62%	23%
Mgt Teams	38%	0%	54%	8%
Clinical Teams	0%	31%	69%	0%

Table 6.2: CS 1 - Ranking of the highest priority for different organisational levels

	Achieving Targets %	Adequate staffing	Patient safety	Achieving financial results
Trust Board	20%	0%	47%	33%
Mgt Teams	13%	0%	80%	7%
Clinical Teams	0%	29%	71%	0%

Table 6.3: CS 2 - Ranking of the highest priority for different organisational levels

While there are some differences in the apportionment of percentages between the two case studies, the patterns of prioritisation exhibit some commonalities. 'Patient Safety' is the highest priority at all levels. 'Achieving financial results' is second highest at the 'Trust Board' level and 'Achieving targets' is the second highest at the 'management team' level. The 'Clinical Team' level results are very similar with around 70% ranking 'Patient safety' highest. Interestingly in both results, 'Adequate staffing' is ranked highest by around 30%, whilst targets and finance do not feature as the highest priority for 'Clinical Teams'. Of note is that 'Adequate staffing' is not the highest priority at the other two levels. The full results are in Appendix 6.3.

The results presented above, whilst limited to the small number of interviewees, shows that 'patient safety' transcends each organisational level as the highest priority issue. The results also show that the interviewees perceive that the Trust Boards identify 'safety' as a top priority area. The Boards are also thought to place considerable importance on the production components (finance and targets). Conversely, at the 'sharp end' (Reason, 1990, Cook and Woods, 1994), the clinical teams particularly are perceived to focus upon the less well defined boundaries of patient safety and staff workload. A possible explanation of the emphasis placed on patient safety, within each of the organisational levels studied, is the Hawthorne Effect. The presence of the researcher investigating patient safety in hospitals will have influenced some of the interviewees to answer the questionnaire placing greater priority to patient safety.

To guard against this bias interviews were subsequently undertaken with each of the staff groups. The interviews explored the actual experience of the staff in relation to themes represented by the boundaries. The data was analysed to draw out inferences about the relative importance of the boundary themes expressed through the prioritisation process. Interestingly, the interview data produced a different result. The results are set out below but in summary the themes relating to finance and targets, in particular, were found to dominate decision making. While patient safety and staffing constructs were relevant, they did not feature in the interview data to the extent anticipated by the questionnaire ranking results above.

In the next section each boundary is discussed separately to identify data relevant to the dimensions of 'visibility', 'movement', 'location' and 'buffer capacity'. It is recognised

that there is interaction between the boundary dimensions, which is also discussed and will be explored in greater depth in later chapters.

6.3 Patient Safety Boundary

The results presented above (Tables 6.2 and 6.3), indicate that patient safety is seen to be regarded as the highest priority at the three different organisational levels in both case study hospitals. An observation from these results, and from the interviews, is the priority given to both staffing levels and patient safety by clinical teams. Many described staffing levels as being crucial to achieving safety for patients.

“...staffing numbers relates to the safety of the patients really that I think you need the patients there, the staff there, to ensure that patient safety is happening really because without the staff you can't monitor that and you can't, if you've got patients climbing out of bed and falling and you've not got enough staff then obviously that's going to potentially happen and those patients are then at risk of you know injuring themselves really.” (Ward Sister CS 1) [CS 1: SWE; Patient safety]

“... the situation on ward 12 is such that patient safety is being put at risk because of [a]shortage of staff. Since ward 12 was designated as the thoracic ward in December 2004 the number of nursing staff for ward 12 has dropped by over a third.” (Nurse CS 3) (Francis QC, 2010a) p.222

(Further coded data about the boundary dimensions is presented in Appendix 6.2)

In contrast, those with wider management responsibilities, after patient safety, were perceived to place a higher priority on finance and targets than staffing. However, when each of the staff groups were asked to explain what constituted patient safety and how they measured it, other than the control of infection, responses were generally vague. This suggests that aspects of patient safety failure boundary were not clearly visible. However, when particular patient safety themes were explored in more detail, a mixed picture emerges as to the 'visibility' dimension of the patient safety boundary.

From observations at the clinical delivery level there were numerous safety checks in place, for example, checking the identity of a patient prior to a procedure. However, these checks were rarely mentioned in interviews. There were limited systematic methods of auditing the reliability of such procedures taking place. The most common reliability audit made was the compliance to hand washing standards by clinical staff [CS 1: Doc 1.5]. From a conceptual perspective, certain aspects of what constituted the

patient safety boundary at an individual patient level, was highly 'visible' to clinical staff. They have been trained in specific procedures that they carry out in order to keep the individual patient from breaching the safety failure boundary. There was clear evidence of clinicians engaging in specific identifiable actions, during patient consultations, which are undertaken to keep patients safe (FN 1-4).

The interview data from both hospitals suggests that the main theme of patient safety was the need to eliminate certain healthcare acquired infections. There was daily monitoring of the number of new Meticillin Resistant Staphylococcus Aureus (MRSA) and Clostridium Difficile (C.Diff) infections [CS 1: Doc 1.6; FN 4]. There were also clear plans detailing how buffer capacity is to be used to isolate and contain any C.Diff or Norovirus (airborne sickness) outbreak [CS 1 Doc 3]. Patient screening for MRSA had become standard practice and new clinical procedures were implemented to reduce the risk of passing on infection [CS 1: Doc 1.5]. The identification of these plans is unsurprising as there are national targets for the reduction of MRSA and C.Diff (Department of Health, 2007, Department of Health, 2008c). (Although there are targets for the reduction of certain infections, as this is a patient safety issue rather than a waiting time target, the data relating to MRSA and C. Diff has been included in the patient safety failure boundary.)

The setting of targets to reduce rates of certain types of infection illustrates the four dimensions of the boundary construct (Figure 6.6). First, the 'visibility' of the boundary is increased with clearly defined types and rates of infection set out as national targets. Second, the 'movement' of the marginal zone boundary inwards as a new norm of acceptability is determined. Third, the result of moving the marginal boundary is that there is an increase in the 'buffer capacity' deployed to reduce the number of those infections. Fourth, the new location of the marginal boundary emerges from a process of public concern, articulated by politicians and detailed in national standards (Department of Health, 2007).

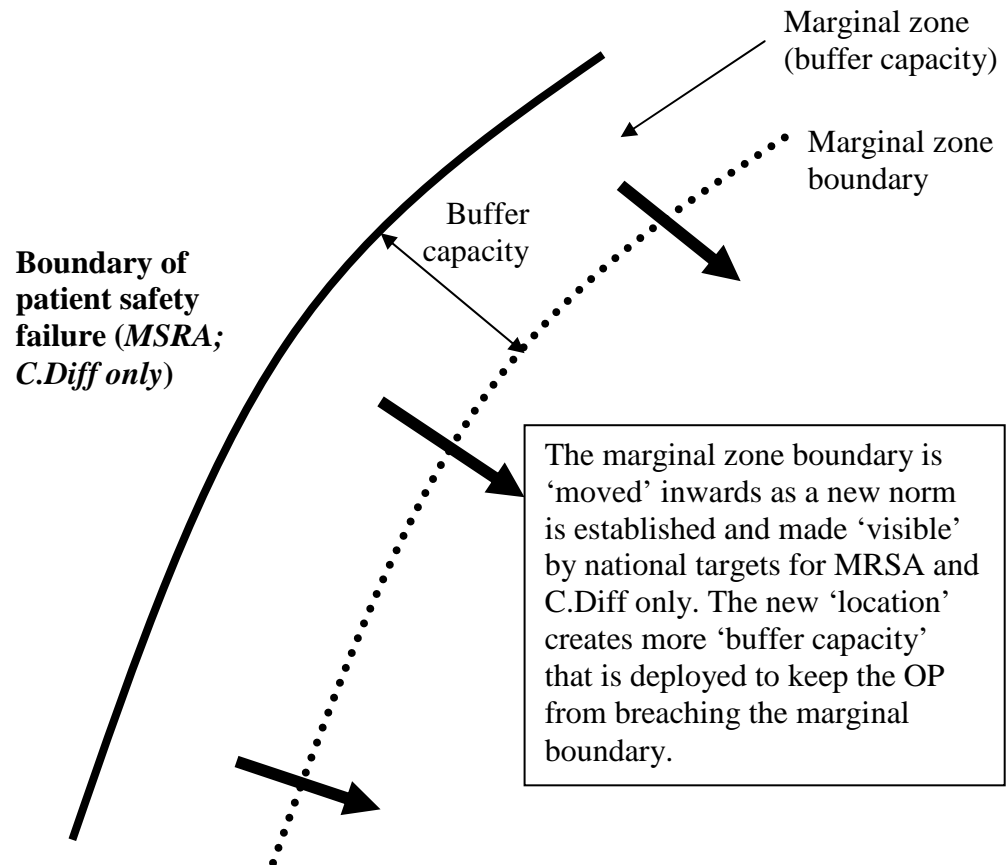


Figure 6.6: Movement of the marginal zone boundary for control of MRSA and C.Diff

However, there was evidence that other types of infection did not receive the same degree of attention, and rates of infection were not routinely monitored [CS 1: Doc 1.5; CS 2: Doc 2.10]. This situation is identified as a common occurrence for the NHS in England (House of Commons Public Accounts Committee, 2009). The CS 1 Annual Infection Control Report for 2008/09 reported a reduction of 17% in MRSA and a 32% reduction in C.Diff on the previous year [CS 1: Doc 1.5]. However, the same report highlighted the multiple ward outbreaks of Norovirus stating that: "Spread of Norovirus, across multiple wards, is exacerbated by high bed occupancy and movement of patients and staff within the hospital setting." This statement did not appear to bring any change to the bed management practices that were observed [FN 2.40-60]. Conceptually, it can be argued that infections other than MRSA and C.Diff did not have the same 'visibility' and therefore did not receive the same degree of attention. However, the measuring and monitoring of certain infections does suggest that parts of the patient safety failure boundary were well defined and therefore 'visible'.

Other measures of patient safety derived from the data are the number and type of incident reports. Patient falls, which create a significant level of harm and mortality in hospitals (Healey et al., 2008) were routinely reported [CS 1.9,13 CS 2.3]. However, from interviews and observations there was evidence of a general level of 'normalisation' amongst nurses and other staff; falls had become 'one of those things' that occur in hospitals [FN 2.28-29] (Grenier-Sennelier et al., 2002, Williams et al., 2009). For those staff who normalise the situation, it can be suggested that the marginal zone boundary is allowed to 'creep' or 'drift' outwards (as seen in point 1 in Figure 6.2). One patient in CS 1 fell more than ten times before a final fall, which proved fatal. In CS 1 action to improve the situation was taken after two fatalities from falls (Williams et al., 2009). Such action to reduce the rate of falls can be viewed as seeking to 'move' the marginal boundary inwards to a position where a fall is considered unacceptable (as seen in point 2 of Figure 6.6). With the increased 'visibility' of the patient safety failure boundary relating to falls, nurses changed their practice. Part of the change in practice was dedicating more nurse time to the prevention of falls. This can be conceived of as creating greater 'buffer capacity' to prevent inpatient falls as the marginal boundary moves away from the failure boundary (Figure 6.7).

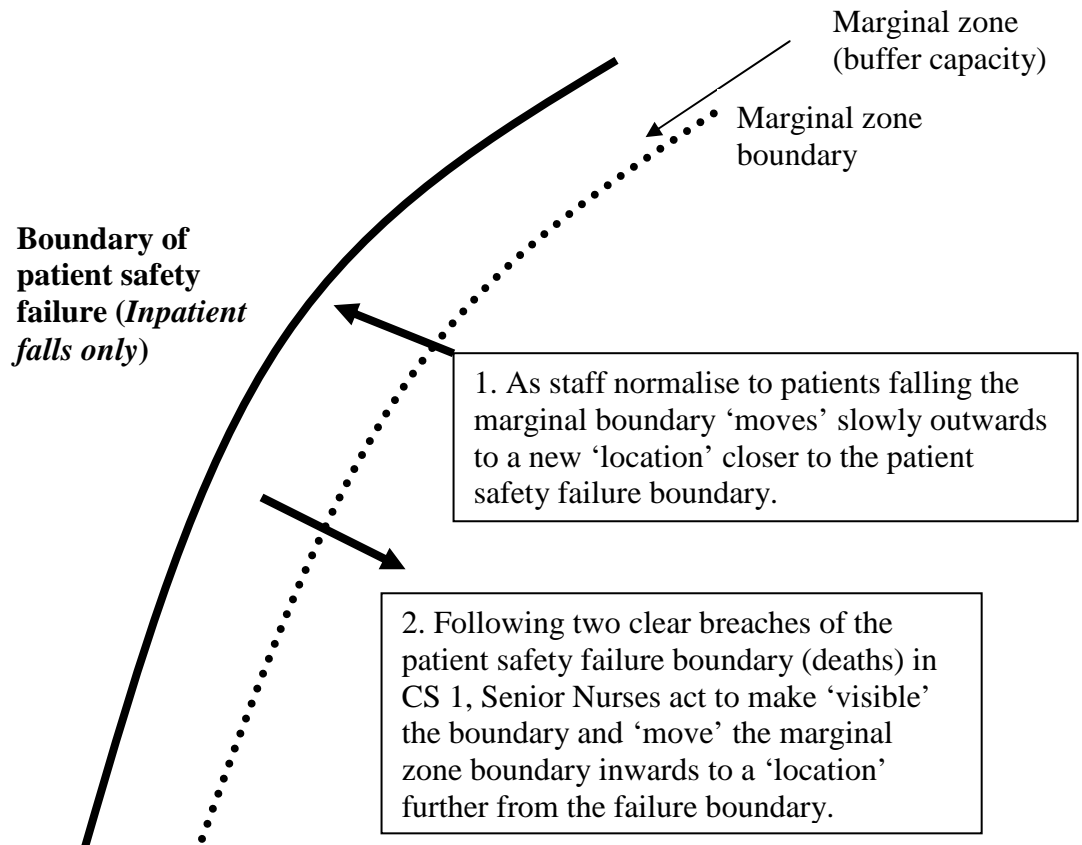


Figure 6.7: Movement of the marginal zone boundary for inpatient patient falls

The only sure way for the safety failure boundary to be 'visible' is when an accident occurs (Cook and Woods, 1994). This is particularly stark when the accident generates an obvious consequence, such as a number of causalities. In healthcare systems, such as hospitals, incidents occur where it is clear that something has gone wrong and where a patient suffers an observable consequent harm or even death. However, the nature of healthcare is that acts of omission or commission that create harm are not always spotted or even counted (Olsen et al., 2008).

Staff were encouraged to report incidents and investigations were conducted where an incident was considered serious. The Boards and Departments in both CS 1 and 2 hospitals received serious incident reports together with analysis of the type of incident and the associated trend [CS 1: Doc 1.9; 1.13; CS 2: Doc 2.13]. The interviewees showed little appreciation of the likely number of incidents that are *not* reported (Sari et al., 2007a, Vincent, 2007, Waring, 2005, Olsen et al., 2008). Therefore, the inference is

made that staff at the higher organisational level assumed that patients are generally being kept safe, as there were few contrary indicators [e.g. CS 1.7; CS 2.5.]

Conceptually, the reported incidents provide ‘visibility’ to some parts of the patient safety failure boundary. However, the research literature clearly indicates that the majority of incidents in hospitals are not reported. Therefore, this ‘visibility’ must be regarded as incomplete, especially to the higher levels in the organisation.

One possible explanation of why many incidents that occur are not spotted or reported is the workload of staff, which is covered in more detail in section 6.3. It is clear from the data that staff regarded the pace of work to be problematic. This created the consequent implications for them to be able to fulfil the needs of patients in a timely and effective manner and therefore guard against potential incidents.

“I would say that staff are so overwhelmed with the number of patients they are actually not clearly focusing on what they need to be focusing on especially if it’s junior members of staff and I think it’s things like drug administration gets given late because they are worrying about getting patients to theatre so some of the routine things might just go aside and it’s not noticed straightaway because the patients would probably not instantly show any signs of suffering or of lack of care but actually over time if you continually don’t give drugs on time it might have a knock on effect to those patients but it’s almost an unseen risk I would say.” (Matron CS1) [CS 1:SWE; Patient safety]

“I guess you know the speed that you are working at, you are working at an awful lot of speed so you are screening and filtering a lot of information quickly and that can incur error. Things are being missed.” (Ward Sister CS1) [CS 1:SWE; Patient safety]

There are a number of interconnected issues surfaced by this data, which suggest that the model of a SWE has to include dynamic interactions between the boundaries. The first is the link between the number of staff and their ability to manage the workload of patients. The workload is influenced by both the waiting time targets and financial requirements placed on the hospital. The second is the unknown impact of delays in drug administration, or incomplete information at the point of patient handover from ward to ward, which illustrates the lack of ‘visibility’ of aspects of the patient safety failure boundary. Therefore, it is difficult to know if harm has occurred or whether the safety boundary has been breached. While a mortality rate, which exceeds expectation, may be used to highlight safety problems, the effectiveness of this indicator is critically debated. For example, the measurement of allegedly high mortality rates is what alerted the Healthcare Commission to investigate the Mid Staffordshire hospital. However, the

subsequent Inquiry, has raised doubts about the usefulness of such a measure (Francis QC, 2010a).

In CS 1 and 2 there was a growing realization that other aspects of safety were important but few if any measures existed for doctors, managers and nurses to monitor patient safety in real time. The routine performance reports (during the study period) to the hospital Board in CS 1 did not include aspects of patient safety, other than the largely process standards required by the Healthcare Commission (now Care Quality Commission) [CS 1: Doc 1.19]. In CS 2 recent changes included the monitoring of a range of quality issues which started to be reported to the Board [CS 2: Doc 2.10]. There was an inbuilt assumption on the part of some board members that patient safety was not a risk issue:

“...you assume that everything is fine because it is run by a senior manager and I am sure she would shove things up if there was a problem. I am on the Risk Committee and patient safety does not figure as a risk because we have this operational team doing things, that’s our assurance you know there’s a team doing this.” (Director CS 2) [CS 2: SWE; Patient safety]

The difficulty in defining, measuring and therefore making ‘visible’ the patient safety failure boundary, both internally to a hospital and externally by policy makers is also evidenced in the case study of Mid Staffordshire NHS Foundation Trust (FT). Monitor, the Independent Regulator of NHS Foundation Trusts, approved Mid Staffordshire NHS Trust to become a FT just months before reportedly high mortality rates at that hospital sparked an investigation. The subsequent investigation found serious shortcomings in the care of patients and in the leadership and governance of the hospital (Healthcare Commission, 2009a). In an internal audit report conducted by KPMG, on behalf of Monitor, they concluded that:

“...whereas it was clear in 2004 what constituted the threshold for quality, it is now less clear what set of factors is regarded as the key data set for evaluating quality performance at an aspiring FT.” (KPMG, 2009) p.4

The situation back in 2004 was that only NHS Trusts who had achieved ‘three star’ status could apply to become an FT (Department of Health, 2002). This approach was replaced in 2006 by the Healthcare Commission’s ‘Annual Health Check’ rating, which included considerable self-assessment and covers a wider range of issues (Healthcare Commission, 2006a). It is debatable as to whether the star rating system was any better at defining what constitutes the quality and safety requirements for patients. The star

rating method looked at a narrower range of target related performance and does not provide detail measures of quality or safety (Department of Health, 2002).

The focus of Monitor in assessing Mid Staffordshire and other aspiring FTs was on the financial and performance management (KPMG, 2009). Monitor apparently believed that other supervising bodies would have flagged concerns around quality and safety (Monitor, 2009). Monitor has since changed the 'Compliance Framework' (Monitor, 2008) to include the requirement for FTs to certify that:

‘...the board is satisfied that it has and will keep in place effective arrangements to monitor and improve the quality of healthcare provided to its patients.’ (p.12)

What is meant by 'quality' is not clearly defined by Monitor. The Darzi Review (Department of Health, 2008b), however, specifies 'quality' as containing three elements. First, patient safety and the concept of doing no harm to patients, second, the patient's experience in terms of compassion, dignity and respect and third, the clinical effectiveness of treatment in terms of outcomes for patients. In 2009/10 there was a requirement placed on NHS Trusts to produce an annual 'Quality Account' covering the three quality elements. However, it should be noted that the content of such reports is left to local determination (Department of Health, 2010e). The initial Quality Accounts published by hospitals have been assessed. The research is limited due to the number of reports examined. However the literature does report many weaknesses, poor definition and measurement of quality, including patient safety (West, 2010, Foot, 2011). This means that there is wide variation in the type and range of measures being used and 'patient safety' appears to remain as the broad concept of 'no harm'.

The subsequent Independent Inquiry into failings in the standards of care at Mid Staffordshire showed that there was widespread concern amongst clinical staff about the level of staffing and the standard of care being provided (Francis QC, 2010a, Francis QC, 2010b). There were poor systems of measurement and governance relating to patient safety and quality at the Trust. It may be suggested that senior leaders were not aware of what constituted the patient safety boundary and therefore did not respond to the chronic and even significant breaches in the ability to achieve acceptable levels of safety (Francis QC, 2010a). Data from the oral evidence given to the Francis Inquiry indicates that the sharp end staff made professional judgements about what ought to constitute safe care for the patients in Mid Staffordshire (Francis QC, 2010b). However,

there did not appear to be a unified view or set of measures that defined the safety failure boundary. The attention of senior managers was on the more ‘visible’ and therefore arguably, to them, the more important boundaries of targets and finance.

While there were measures used to judge NHS organisations performance on patient safety (Healthcare Commission, 2008), it is surprising to note that no reference was made to them during the data collection interviews. An examination of the Healthcare Commission measures reveal that they are almost entirely process related. Checks are made to ensure that hospitals have processes in place to monitor incidents and to comply with policy and ‘safety alert’ recommendations on issues such as child protection and the use of medical devices. The only direct patient care criteria is about MRSA infection.

To summarise, this data shows that the definition of what constitutes patient safety is often lacking or weak, both within healthcare organisations and across external agencies. Therefore, apart from some elements of patient safety such as certain infections, patient falls and reported incidents described above, there was a lack of ‘visibility’ of the patient safety failure boundary and the associated marginal boundary. For the elements of patient safety that are ‘visible’, then it is possible to conceptualise the boundary ‘movement’, ‘location’ and ‘buffer capacity’. The examples described are summarised in Table 6.4. Where there was a lack of ‘visibility’, as evidenced from the CS 3 data, the boundary of patient safety failure is vulnerable to being breached.

Construct	Dimensions	Examples from the data
Patient safety failure boundary	Visibility	National targets for MRSA and C.Diff infections
Patient safety failure boundary	Movement	Staff normalise to patient falls moved the marginal boundary outwards.
Patient safety failure boundary	Location	Senior nurses state that ‘no fall is acceptable’ which moved the marginal boundary inwards and providing a clear ‘location’ marker (no falls).
Patient safety failure boundary	Marginal zone (buffer capacity)	Introduction of the screening of patients for infections and identification of resources to contain the spread of infection.

Table 6.4: Dimensions of the SWE ‘Patient safety failure boundary’ construct with examples from the data

There has been a growing international effort to ensure that patient safety is clearly defined and managed as the scale of unintentional harm has become more widely known (Vincent, 2010). Patient safety campaigns, together with reports on failures in healthcare provision, have highlighted the problems inherent with a leadership focus on the productivity agenda (Healthcare Commission, 2006b). Following the Darzi Review (Department of Health, 2008b) a greater emphasis has been placed on all aspects of quality within the NHS. The publication of the Inquiry into Mid Staffordshire has also increased the development of measures of safety and the wider quality agenda. Interestingly, the evidence detailed below from the case studies shows that the other boundaries of finance, targets and workload tend to be more clearly visible. This has further implications for patient safety which will be explored in more detail in Chapter 9.

6.4 Unacceptable Workload Boundary

The SWE model depicts a boundary relating to an unacceptable workload for staff. Individuals and groups have different capacities to do work, so defining what is acceptable is difficult. However, most organisations seek to make some form of capacity planning to identify the number, type, and skills of staff they need to meet the workload. NHS hospitals have many professional and other staff that form the team that care for patients.

Data from CS 1 and 2 shows that managers plan the number of staff expected to work in clinical areas and in support functions [FN 1.54; CS 2.1]. Budgets were set according to the numbers and seniority of staff employed in each area. The overall staffing levels and day to day fluctuations were closely monitored and controlled by senior nurses, doctors and other managers. Where shortages of nurses occurred, for example due to sickness, then decisions were made about redeploying staff.

“We have within our division we have a staffing template which we know how many staff we should have on each shift and we know on a day to day basis if the staffing drops you know how we can, you know whether that’s acceptable or not acceptable but we also have to think about what’s actually happening on the ward at the time so we’ll share staff around to make sure that every area is covered as you know as well as we can.” (Matron CS 2) [CS 2: SWE Staff workload]

Given the close attention to nurse staffing in particular, it can be argued from a conceptual perspective that for front line staff, the unacceptable workload boundary was 'visible' and therefore clearly 'located'. There was some flexibility to manage the situation. The 'buffer capacity' was created through the ability to call in extra nurses from the Nurse Bank or Private Nursing Agencies to fill gaps in the rota. However, at busy times or in holiday seasons, nurses were not always available in the required numbers [FN 1.78]. Nurse interviewees, in particular, made the link between the staffing numbers, the seriousness of the patients' illness and safety. The senior nurses use their professional judgement to move personnel around to cover potential gaps in staffing [CS 1: Doc 1.20].

There was less flexibility with junior medical staff as finding locum staff was often not possible. For example, during an outbreak of the Norovirus sickness bug some staff took sick leave, leaving wards with minimal doctor cover.

“But junior staff wise, I don't you just have to kind of grin and bear it essentially and at that time the staffing was very thin and I don't know if that would have jeopardised patients' safety a great deal, it meant that there were a whole lot of wards with one doctor on it.” (Doctor CS 1) [CS 1: SWE; Staff workload]

Part of the context for the staffing of junior doctors is the European Working Time Directive, which has reduced the working hours of junior doctors (Department of Health, 2009a). The regulations result in junior hospital doctors working in shift patterns with a consequent reduction in continuity of care for patients and reduced team working with the same consultant.

“Well I think that there is a huge problem with financial pressures, target pressures and also with the European Working Time Directive that is removing the continuity of care from patients which I think is highly detrimental to quality care and I've seen it personally, I've seen it with colleagues, I've seen it with my daughter where a failure of continuity of care has led to potentially a dangerous situation arising.” (Doctor CS 1) [CS 1: SWE; Staff workload]

Changes in the training of junior doctors, the increase in the number of female doctors and the restrictions on the number of non European Union doctors, has meant that filling gaps in the rotas is often not possible [CS 1.25; FN 3.10]. When a hospital faces pressure to achieve targets or meet peaks in demand then the adaptive capacity of the key medical resource was stretched.

“One of the things is that everything is so target driven everybody is working extremely hard so money is being thrown at the system, consultants are being paid high rates to do additional clinical sessions and basically time shift there their professional development work into the evenings and weekends which they either do or don’t do but you know it’s their choice. As a result of that they’ve got no capacity to step up and do extra work, so I’ve got two extra outlie wards I’ve got to look after. So we’ve come down from 110 medical outliers to 60 but we are running the whole of ward 10 which is vascular surgery and the whole of ward 19 which is orthopaedics as medical wards. So suddenly I’ve got to find a whole medical team to operate those wards so I am having to spread them thinly from elsewhere. At the same time because I can’t fill posts I’ve got 8 vacancies okay some of which is driven by pregnancy so that’s another cause for less resilience in the system is a lot of the junior doctors are female and therefore we’ve got a high pregnancy rate.” (Doctor CS1) [CS 1: SWE; Staff workload]

Conceptually, the marginal zone boundary for doctors was closer to the unacceptable workload failure boundary than for nurses (Figure 6.8). The reason for the different locations of the marginal boundary was the available ‘buffer capacity’ for the nursing staff, which was not present for medical staff. This leads to the view that there are multiple marginal boundaries to be considered for different staff groups.

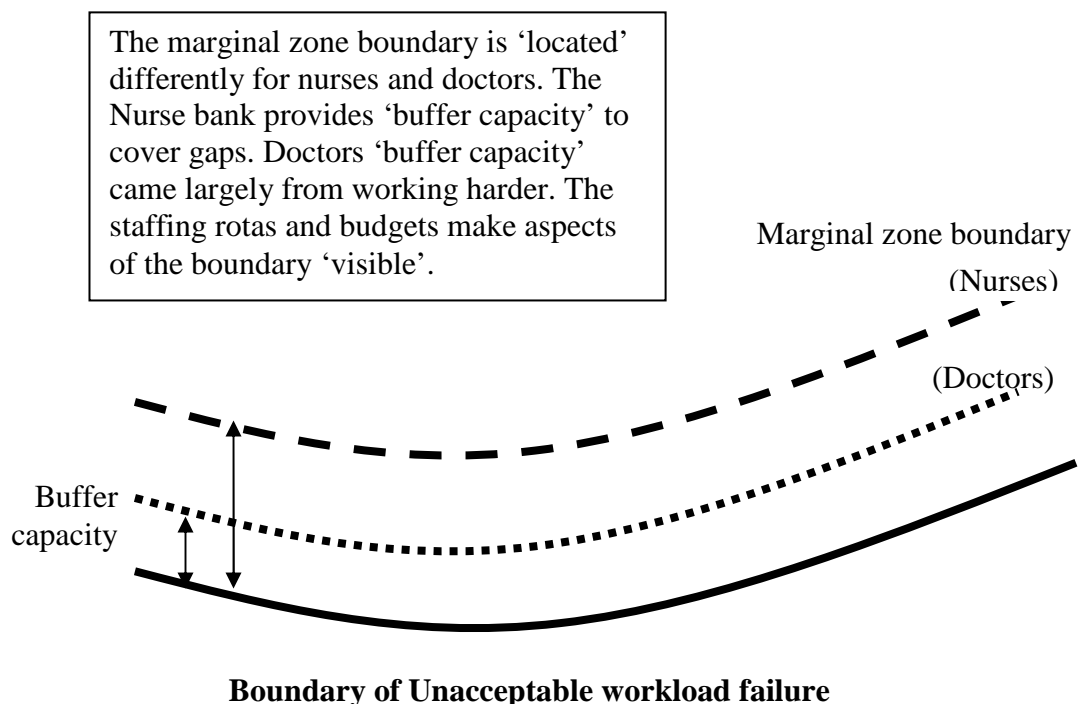


Figure 6.8: Location of the marginal zone boundary for nurses and doctors

There was evidence that the sharp end nurses and doctors knew what constituted an unacceptable workload. An example of this were the comments made about the

proposed changes to staffing levels by one senior nurse from CS 3:

“... the impression I had was on some wards we would have three trained nurses trying to cover effectively two wards and probably a couple of untrained support staff. In my estimation that is nowhere near enough”. (Advanced Nurse Practitioner) (Francis QC, 2010a) p.211

Often the limit of what was acceptable is stretched and efforts are made to support staff and provide replacements to fill gaps when that is possible. Senior managers in CS 1 and 2 recognised that at busy times, the staff put in considerable extra effort [CS 1.4; 1.15; CS 2.1]. Managers took action to ensure that there were no delays in recruitment to vacant nursing posts [FN 1.78]. In CS 1 the hospital Board invested money in new electronic beds and other equipment to improve the working conditions for staff.

“...so we invested £5 million last year in terms of revenue, very much targeted at the clinical teams coming up with the things that would make life easier.” (Director CS 1) [CS 1: SWE; Workload]

However, evidence from CS 3 shows that key decisions were taken to reconfigure wards and reduce nurse staffing levels (Francis QC, 2010a). Managers making those decisions did not appear to have listened to or acted upon the concerns of front line staff. The changes appear to have been largely motivated by the need to save money (Healthcare Commission, 2009a).

“We noted from the figures for staff in post that the largest reduction in the number of nurses occurred between April 2006 and April 2007. The reduction of staff in post in that 12-month period was nearly 130 whole-time equivalent nurses.” (Healthcare Commission, 2009a)

The dominant context for the managers in CS 3 at that time was to achieve FT status. As noted above, much of the assessment for FT status was related to the financial and operational management in achieving targets (KPMG, 2009). As the staff workload boundary is often a matter of professional judgement, it is less well defined and therefore, not as ‘visible’, to managers as the financial and target boundaries. In CS 3 decision makers did not appear to make the balanced judgements to the same extent as those found in CS 1 and 2.

However, in CS 1 and 2 there was little if any knowledge expressed about the human

factors implications for patient safety of high workloads, staff working long hours or in unfamiliar surroundings (Reason, 1990). For example, there is evidence in the literature that nurses and doctors who work long hours are much more likely to make mistakes (Rogers et al., 2004, Scott et al., 2006, Fahrenkopf et al., 2008). The working hours of junior doctors were monitored through a self reporting scheme which was open to under reporting the actual hours worked. The hours worked by other groups of staff was not measured. They were assumed to have worked the hours on the rota. Observation and interview data confirmed that many nurses worked longer and often missed breaks.

“And they work you know but they are under pressure all the time, there are days where they, I mean and now they are just stopping for lunch.

And it's now

Twenty past three.” (Ward Sister CS 1) [CS 1: SWE; Staff workload]

In conceptual terms the unacceptable workload boundary is more visible closer to the sharp end of providing patient care. This was evidenced by the questionnaire results, interview and observational data. Nurse managers in particular know what levels of staffing should be in place [CS 1.15; 1.18; CS 2.7; 2.11]. Both doctors and nurses make judgements about how to use the staffing resources flexibly when required by changing circumstances. Senior managers often recognised that staff worked harder during peaks in demand but did not appear to appreciate the potential patient safety implications. That means that the marginal zone boundary can ‘drift’ outwards as staff and managers ‘normalise’ to working harder.

It is concluded that the marginal zone boundary related to the unacceptable workload failure boundary was not set in a fixed location either by the staff themselves or by the senior managers. The relationship of the unacceptable workload boundary to the patient safety failure boundary is conceptually more visible to front line clinical staff. Managers appeared to understand the link between staffing numbers and costs very well. They also made judgements about staffing levels and the ability to meet targets [e.g. CS 2.1].

Whilst professional bodies make recommendations about what the staffing levels should be (Scott, 2003), it is apparent that the staffing levels are often negotiated internally. Planned changes to staffing levels were agreed through the business planning process. At times of perturbation the internal professional judgement of staff provided the

‘visibility’ for senior managers who sanctioned additional expenditure, on for example, agency nurses. Evidence from observation showed that senior managers in CS 1 and 2 listened carefully to staff and made judgements about what was reasonable workload pressure. However, there were times when circumstances were beyond their control and the staff workload escalated.

CS 3 is an example of where the judgement about staffing numbers and workload was influenced by the context of financial control. It may be argued that in CS 3 the reduction in nurse staffing levels ‘moved’ the marginal zone boundary outwards to a ‘location’ close to the failure boundary, and therefore reduced the ‘buffer capacity’. The lack of ‘visibility’ of the unacceptable workload boundary in comparison to the financial and target failure boundaries is a likely conceptual explanation for the reduction in staffing levels.

There are four conclusions to be drawn from this data:

- First, that the ‘visibility’ of the unacceptable workload boundary and associated marginal zone boundary was greater for staff working close to providing direct patient care. The interrelationship between workload and patient safety was also more ‘visible’ to front line staff.
- Second, that staff and managers normalised to having to frequently work harder, thus the marginal zone boundary ‘moved’ slowly outwards.
- Third, that there were different marginal zone boundary ‘locations’ for doctors and nurses, due to the difference in the availability of ‘buffer capacity’.
- Fourth, the interrelationship of staffing levels with finance and targets was more ‘visible’ to managers and directors than the relationship of staff workload to patient safety. This is demonstrated in the next section.

Examples of the dimensions of the unacceptable workload boundary are summarised in Table 6.5:

Construct	Dimensions	Examples from the data
Unacceptable workload boundary	Visibility	Staffing rotas and budgets provided the basis for what is agreed as acceptable staffing levels.
Unacceptable workload boundary	Movement	As staff normalised to working harder during busy periods the marginal boundary ‘drifts’ outwards.
Unacceptable workload boundary	Location	The location of the marginal zone boundary was different for doctors and nurses.
Unacceptable workload boundary	Marginal zone (buffer capacity)	The buffer capacity for doctors was largely created by them working harder / longer (“I am having to spread them thinly from elsewhere”) Nurses have the Nurse Bank and Agency Nurses to provided some buffer capacity.

Table 6.5: Dimensions of the SWE ‘Unacceptable workload boundary’ construct with examples from the data

6.5 Financial Failure Boundary

Despite record levels of spending on the NHS, there was a period in the mid 2000s when financial control in some organisations was weak leading to an overall overspend for the NHS in England (House of Commons Health Committee, 2006). The situation led to a renewed emphasis in 2006/07 by the Department of Health on keeping organisations within budget and where possible to make a surplus (Department of Health, 2006a). There are nationally set tariffs for clinical procedures and each hospital agrees a contract with their local Primary Care Trust (PCT). That contract forms the basis for the financial budget each year.

The case study hospitals have different histories of financial management. CS 1 was a ‘first wave’ FT with a strong record of achieving financial plans. CS 2 has a history of large financial overspends. Over the past three years the overspending has been reversed and budgets balanced. CS 3 faced some financial challenges and therefore sought to make large scale savings as evidenced from one of the inquiry reports:

“The year 2006/07 was a challenging one for the NHS, as trusts were required to achieve financial stability. That year, the trust set itself a challenging agenda to meet national targets for cost improvement, stabilise its finances, and become an NHS foundation trust. The trust set a target of saving £10 million, including a

planned surplus of £1 million. This equated to about 8% of turnover. To achieve this, over 150 posts were lost.” (Healthcare Commission, 2009a)

Financial management deals in numbers. Providing that there are reasonable processes in place to set budgets, attribute income and costs, then the defining and monitoring of the financial failure boundary is relatively straightforward. Therefore, the financial failure boundary is ‘visible’ to those with budget responsibilities. Trust Board reports for CS 1 and 2 show very detailed monthly financial reports. In CS 2, due to the history of overspending, all the non-executive board members were part of the Finance Committee of the Board which scrutinise the monthly reports.

“Yes they are detailed and they are, we have a very lengthy finance committee meeting and go through the whole thing in great detail and at the Board it’s virtually a rubber stamp situation because we’ve gone through it all before.” (Director CS 2) [CS 2: SWE; Finance]

There were clear processes of monthly, quarterly and annual reviews within the CS 1 and 2 hospitals to monitor the financial position [CS 1: Doc 1.1; 1.2; 1.8; 1.9; CS 2: Doc 2.2; 2.2; 2.14] . At a senior management level there were resources and processes in place to manage and monitor the financial situation including individual ward and department level.

“It’s very clear we get the monthly income and expenditure reports that we’ve always had in exactly the same format as they have always been and also we have a very robust directorate and management accountant set up and probably the most robust that I have ever known is as much that we have a Divisional Accountant that’s a member of our team and very much dedicated to what we do so they get to understand what we do so that they don’t just record and report they advise into the whole set up...” (Senior Manager CS 1) [CS 1: SWE; Finance]

In this case the boundary is highly ‘visible’ and therefore easy to ‘locate’ at different levels in the hospitals. As noted above, managers were very aware of the interrelationships between staffing and costs as over sixty percent of the budget was spent on pay. Linked to the boundary of unacceptable workload, managers balanced the financial implications of filling vacancies and covering sickness on the wards with the need for strong financial control [CS 1.18; CS 2.11]. Senior nurses and ward managers reported using their budget allocation flexibly. However, it is also clear that they had to account for any overspend and manage within their overall budget [CS 1: Doc 1.8; 1.9].

This next quote is an example of where the staffing is considered primarily from the financial perspective by senior managers.

“On a more operational level we run our monthly performance reviews which is our opportunity to sit down with the divisional teams and look at their monthly I and E position and the cumulative year to date and their year-end forecast position and we will have a number of not just the budget information but we would have a number of indicators that we would look at around agency staffing usage, vacancy control information and no vacancies get approved for filling without an exec team, they all come to the exec team every week and similarly agency usage is only approved by myself or the Director of Nursing and out of hours on call etc. We have, we regularly review short and long term sickness levels and those sorts of things as indicators to underpin what’s going on with the pay position, the variable pay position.” (Director CS 2) [CS 2: SWE; Finance]

The data indicates that CS 1 and 2 had strong financial management processes. They set a financial plan for the year and managed the within year variations as required. At times of high demand for services (or perturbation), financial considerations were not at the forefront of the decision making process. Nurses, doctors, managers and directors, in both case studies, made the similar point that in extreme situations staffing or safety considerations outweighed the financial consequences:

“...whenever you are kind of up against it we never, we never make a decision based on let’s not cover a shift because it will save money and we would never not get agency in if we really felt it was needed for safety or get security if a patient is quite violent. So we would always, so money is not part of the equation in terms of safety. Not in the context you describe which is when the hospital is busy.” (Senior Manager CS1) [CS 1: SWE; Finance]

Conceptually, there was some ‘buffer capacity’ used to meet the immediate costs in times of perturbation. Some flexibility was given to short term expenditure to keep the other boundaries from being breached. For example, during times of staff sickness or exceptional demand for inpatient beds, additional resources were spent to bring in bank and agency nurses. Additional money was paid to consultant medical staff to achieve targets. Long term increases in staffing required a full business case to argue for the additional expenditure, which then took time to be considered. There was a recognition that a balance had to be struck between the level of resources and safety.

“...there’s no doubt about it if you had lots more nurses and you know more time and you didn’t have bed problems there would be fewer mistakes but it all costs money of course.” (Senior Manager CS1) [CS 1: SWE; Finance]

Short term safety issues, for example, staffing wards to cover gaps in the rota were funded. The short term decisions were accounted for later in performance management meetings. When longer term expenditure was considered to improve staffing levels and hence patient safety, it was often the financial considerations that took precedence. The concept of ‘bounded rationality’ helps to explain that decision makers are influenced by the immediacy and ‘visibility’ of the current needs when making choices (Sterman, 2000, Booth Sweeney and Sterman, 2007). Future needs are weighed up using the available information on projected financial costs. The costs are easier to define and make ‘visible’ than the less well defined safety benefit. The investigation into the governance of the CS 3 Trust found that over a number of years the Board had focused on the more measurable aspects of finance and targets in the context of seeking FT status:

“We analysed the minutes of the trust's board meetings from April 2005 to 2008. The minutes indicated that discussion at the board was dominated by finance, targets and achieving foundation trust status.” (Healthcare Commission, 2009a)

The financial failure boundary was clearly defined and monitored in the CS 1 and 2 hospitals at all levels of budget management from the Board to individual budget holders. Conceptually the boundary could be regarded as highly ‘visible’ across different levels in the hospital system. Financial management is largely judged by the end of year performance. The timing of the judgement allows for some ‘buffer capacity’ through the year. Overspends early in a financial year can be balanced by subsequent underspends [FN 1.48], thus borrowing ‘buffer capacity’ from the future potential underspends. In CS 2 the financial ‘buffer capacity’ was borrowed from under spending departments [CS 2.13]. The financial ‘buffer capacity’ borrowed from the future had to be paid back. This can be regarded as allowing the marginal zone boundary to move outwards with the plan to reinstate it to the original position later in the year. This possibility of borrowing from the future diminished close to the end of the financial year as the ability to find savings in the future runs out of time. Therefore, the boundary can be conceived as becoming more and more ‘visible’ and the marginal zone boundary ‘location’ becoming ‘unmoveable’ towards the end of the financial year.

CS 1 had a financial plan to make a budget surplus. When the number of patients treated was higher than contracted for, additional income was generated adding to the planned

surplus [FN 1.58; CS 1: Board Report 26.11.08]. In terms of the SWE, the additional income 'moves' the 'location' of the marginal zone boundary inwards, thereby providing an increased financial 'buffer capacity'. During perturbations there was evidence of some relaxation of the budget constraints on staffing. This situation was not planned. From a conceptual perspective the surplus of income provided the financial 'buffer capacity' to cover the costs of unexpected levels of activity.

There are three conclusions drawn from the data:

- First, the financial failure and associated marginal zone boundary were 'visible' with nationally set tariffs, contracts, budgets, monthly reports and performance reviews.
- Second, the need for strong financial control dominated decision making except in extreme situations where the visibility of the immediate needs, for example nurse staffing, took precedence.
- Third, the financial 'buffer capacity' was achieved from a number of sources such as additional income, borrowing from future underspends, or from under spending departments. There was some 'movement' in the marginal zone boundary which becomes restricted towards the end of the financial year.

Some examples of the dimensions for the financial failure boundary are presented in Table 6.6.

Construct	Dimensions	Examples from the data
Financial failure boundary	Visibility	Budgets, regular reports and performance reviews made the boundary very 'visible' to budget holders
Financial failure boundary	Movement	There was limited movement inwards with additional income; outwards when borrowing from future under spends. Movement diminished as financial year progresses.
Financial failure boundary	Location	Location was set by budgets and subsequent decisions.
Financial failure boundary	Marginal zone (buffer capacity)	Expanded through net income above budget. Reduced through higher net expenditure than budget. Borrowed from future under spends.

Table 6.6: Dimensions of the SWE 'Financial failure boundary' construct with examples from the data

6.6 Target Failure Boundary

Since the NHS Plan was published in 2000 there are a number of targets that hospitals have been required to meet (Department of Health, 2000a). As noted earlier, the NHS 'Operating Framework' documents place a strong emphasis on the reduction of time patients wait to access hospital services. This section outlines the targets that focus on access to inpatient beds. It describes the priority of attention paid to the targets, how they are monitored and the attitudes of staff towards them. It is argued that the boundary of target failure was the most clearly 'visible' 'located', least 'moveable' with the smallest planned 'buffer capacity', of all the four boundaries. This was largely due to the ease with which success and failure of achieving targets was measured and rewarded.

At the time of the research (2008-10), the main targets that impact on hospital bed capacity were:

- A maximum of waiting time of one month from diagnosis to treatment for all cancers.
- Waiting time in Emergency Department (ED) – maximum of 4 hours for 98% of patients

- Waiting time from Referral To Treatment (RTT) – maximum of 18 weeks with a local ambition from the Health Authority to reduce to 15 then 13 weeks. (This target relates to non-emergency patients)
- Cancellation on day of admission rate of less than 0.8% (non-emergency patients)
- All patients to be treated in single sex accommodation

Whilst these targets may appear to be simple, some of them have complicated technical definitions and measuring requirements. For example, with the 18 week target there were definitional complications about what constitutes ‘treatment’ and therefore knowing when the clock stops for a patient in a pathway of care (Department of Health, 2010a, Department of Health, 2010b). Even senior managers, who spent a lot of their time performance managing their organisations to meet the targets, confessed that rules being applied are not always clear or helpful.

“As an organisation I think, the complexity of targets is something that I think is extremely difficult in the health service so I would say that the organisation knows the headline targets so the waiting and access targets. Some of the underpinning targets I think are shrouded in mystery to, not only to them but to me at times so the 18 week is a real good example of that where we are achieving the 18 week target but we are failing on data completeness because the algorithm that they are using is a flawed algorithm. So you know you can fail something even though you are doing the right things.” (Director CS 1) [CS 1: SWE; Targets]

Despite the complicated nature of the some of the definitional issues, targets in CS 1 and 2 were highly ‘visible’ and received a high level of attention, particularly by managers. Observations show a hierarchy of importance linked to the immediacy of the target on the operational management of the hospital. The shorter the target timeframe, the higher that target came in the hierarchy of importance for managers. The 18 week target allowed clinicians and managers time to plan the workload and capacity. There was some flexibility and therefore limited ‘buffer capacity’ within which to manage the timing of patient admissions. With the cancer waiting time target the timescale was much shorter so the opportunity to be flexible was more limited.

The ED target provided a very short time period within which to succeed or fail. Therefore, considerable attention was required to avoid unnecessary delays for the patients. The monitoring of performance against the four hour ED target was continuous

both within the department and by managers elsewhere in the hospital. Patients who were not admitted or discharged within four hours are known as ‘breaches’. The hierarchy of urgency is demonstrated by the frequency of reports given to senior managers, as illustrated by the following quote:

“I am fairly obsessive about that [laughter] I get daily reports on some of the stuff. So I get real time information on, four hour target, I get real time information of any breaches in particular, I get weekly information on where we are with our RTT position” (Director CS2) [CS 2: SWE; Targets]

All the targets are ‘visible’ to managers through daily and more frequent reports. What the data demonstrates is that the ‘visibility’ is linked to the timescale and apparent hierarchy of importance of the target; the shorter the timescale the more visibility is required. For example, in CS 1 there was a large computer screen in the bed management office that was linked to the ED computer providing live performance data for each patient on the four hour target. The bed managers sought to manage the flow of patients through the hospital to ensure that there were empty beds to accommodate patients from the ED and other admission routes. Much of their time was spent finding empty beds to allow the ED to meet the four hour target.

“...the Emergency Department position is highlighted every day and is part of this operational forecast you know and on a weekly basis, well on a daily basis we are looking at where we are against the kind of you know the position for that week so very much saying we need to be achieving 98% every week here. So that’s, and this forecast is circulated widely across the organisation every day. At our 12 o’clock bed meeting we review all breaches for the previous day and try and identify any themes so that’s managed very much on a daily basis.” (Senior Manager CS1) [CS 1: SWE; Targets]

The bed managers monitored the ED position in real time [FN 2.48]. Senior managers reviewed the situation regularly. When the ED became busier and the threat to the four hour target more imminent, then the situation was escalated and senior managers became actively involved in managing the bed capacity (FN 2.60). The Chief Executive and other Directors were observed in CS 1 moving patients from the ED to avoid breaches [FN 2.60]. Directors in CS 2 reported taking direct interventions to improve the performance against the ED four hour target [CS 2.2]. The variety of actions taken to achieve the targets is explored in more detail in the next two chapters.

Clinical and managerial staff expressed the view that the four hour target was necessary

to ensure timely and appropriate care for patients.

“I know people might get obsessed about sort of you know breach times and four hour waits and it’s just a target but sometimes the time by which they are seen is a surrogate marker of quality...” (Doctor CS1) [CS 1: SWE; Targets]

There was also strong performance management of the ED target in both CS hospitals. For example, the situation where the achievement against target was just under the required 98% could generate considerable management effort in taking corrective action particularly where there was no ‘buffer capacity’ available:

“At the moment we are below 98% so we are working.

How far below?

When I looked just now it was 97.2% or something but still that does not give us a lot to year end. I’ve met with the A&E consultants, I’ve met with the executive team because they are at the moment conflicting well not conflicting but there are financial concerns and A&E target concerns and we have discussed with the execs today neither of which we can afford to miss out on.” (Senior Manager CS 2) [CS 2: SWE; Targets]

The interview data points to the interaction of the financial and target boundaries which have to be managed. The data also demonstrates the level of ‘visibility’ given to those boundaries. This level of internal performance management is a reflection of the attention placed on the performance against targets by the Department of Health. Progress on targets was closely measured and monitored at the regional and national levels in the NHS (Department of Health, 2008a). During the winter months it was normal for non FT hospitals to have to make daily returns of performance achievement to the Department of Health, via the local Strategic Health Authority [CS 2.1]. Failure to achieve key access targets brought significant attention from either Monitor, as the independent regulator of FTs, or the local Strategic Health Authority [CS 1.6; 1.13; 2.1]. A high degree of importance was therefore attached to achieving the targets by both clinicians and managers. These quotes point to the level of ‘visibility’ given within the hospitals to this boundary:

“...at Christmas time we were teetering on the very edge of kind of target failure which for this Trust is the kind of ultimate a most heinous crime if you like...” (Doctor CS 1) [CS 1: SWE; Targets]

“...we always deliver every target.” (Director CS 2) [CS 2: SWE; Targets]

“(the ED target) is a real priority for the Trust because to fail to meet the target will have significant implications for the Trust and some of the repercussions will have a real affect on the quality of patient care particularly where there are financial benefits involved and so I think the nursing staff particularly see the target as something they very badly need and want to achieve.” (Doctor CS 2) [CS 2: SWE; Targets]

The Independent Inquiry into the Mid Staffordshire NHS Foundation Trust explored the implications of failing to meet targets. The following extract from the Inquiry report shows the level of performance monitoring that was in place externally. Arguably, this can be viewed as the level of ‘visibility’ given to this boundary.

“As will be seen, the Trust Board placed a high priority on compliance with nationally set targets, and, in particular, the four-hour waiting time target for A&E. The pressure to comply with such targets came from the Department of Health (DH), the strategic health authorities (SHAs) and the primary care trusts (PCTs), as explained by the then Chief Operating Officer:

Q: But the consequence of failing to meet a target was essentially that it would reflect poorly on the Trust when compared with others?

A: Yes but it would be more than that because it would be performance managed via both the PCT and the SHA at the time against what was happening, why the required standard wasn't being met and what actions the organisation would take to improve and reach the required –

Q: So your successor in the post at the SHA would be on the phone to you saying: why isn't this target being met?

A: Yes, and what are you doing about it, and similarly from within the PCT.

Q: Was there pressure being brought to bear on you not only from within the Trust but also from the PCT and the SHA to ensure that these targets were met, or to explain why, if they were not met, that was?

A: Yes, and beyond that because ... there was a team within the Department of Health which likewise was looking at any outlier performance and would expect through the SHA an understanding of what was happening and why it wasn't being – why improvements weren't being seen.

Q. And with specific reference to the A&E target:

I think we were all put under pressure to meet the four-hour target. It wasn't just something that was unique to Mid-Staffs. And there was very much a sense from the SHA, the PCT, Monitor, the Department of Health, that that was a required standard that patients should be able to be clinically dealt with within the department within the four-hour threshold. I do not believe anyone used bullying tactics.”

(Francis QC, 2010a) pp.162-3

The pressure to meet targets at Mid Staffordshire was passed down the line to the

doctors and nurses working in the ED. From the SWE perspective, the highly ‘visible’ and non-‘movable’ marginal zone boundary is due to the potential consequences of breaching the boundary to the hospital’s ambitions to become an FT. A further extract from the report written by the Inquiry Chairman, Robert Francis QC, illustrates the attention given to the target failure boundary:

61. An emergency physician told me:

The nurses would go into that meeting and they were told in the meeting that [if] there were any breaches to – that is breaches of the four-hour rule – they would be in danger of losing their jobs. On a regular basis, and I mean a number of times per week, when I was on day shifts, I would see nurses coming out of that meeting crying.

62. The A&E consultant agreed that senior nurses would pressurise junior doctors to discharge patients to meet the target:

[They would] say: look, come on, someone is going to breach in 10 minutes, and sometimes they would be asking the senior to go and sort out the mess or make a decision.

(Francis QC, 2010a) p.165.

There was no evidence of such pressure being applied in CS 1 and 2. However, the very limited or non-existent ‘buffer capacity’ and high ‘visibility’ of the ED target demonstrates the degree to which the boundary of target failure was monitored. The high level of monitoring was present for all the targets both informally and through the formal reporting mechanisms of the hospitals. There was a high level of engagement in monitoring achievement against the range of targets and attention by managers on actions required to resolve actual or potential failure.

“it’s monitored monthly by the Board but you know sort of weekly and daily by other key members, divisional managers, Director of Ops, the Chief Executive, myself you know we are sort of keeping an eye on what’s happening throughout the month and then the Board formally monitors it on a monthly basis.”
(Director CS 1) [CS 1: SWE; Targets]

As noted above, the success or failure of not breaching the financial boundary was ultimately judged at the end of the year. Therefore, there was some flexibility to manage financial overspend and underspends across departments or over time. With waiting time targets that flexibility was much more limited due to the way failure was measured. The lack of flexibility provides an insight into the ‘buffer capacity’ related to this

boundary of the SWE:

“I think it would be fair to say that targets are the most closely managed because you only have to have one or two kind of failures with a target and you can't take them back whereas with your finances if you overspend in month one you can claw it back over the rest of the year. If you have a breach of some description it sits there. So targets are monitored all the time literally every day I will be looking at a range of target indicators.” (Senior Manager CS 1) [CS 1: SWE; Targets]

The different targets created a set of competing requirements to admit patients to hospital within set time limits. There was a considerable management effort to understand the level of demand for access to the case study hospitals. In CS 1 and 2, historical data was used to help predict the number of emergency admissions [CS 1: Doc 6; xls 1.1; CS 2: xls 2.2]. The number of patients needing planned admission was constantly monitored and communicated to decision makers. The pattern of discharges from the hospitals was also monitored. Managers sought to understand what actions were required to meet the demand for treatment within the waiting time targets. This can be conceptualised as keeping the OP away from breaching the target failure boundary.

In summary there are four conclusions to be drawn about the target failure boundary:

- First, the target failure boundary was highly ‘visible’ with a series of nationally set targets combined with strong performance management at all levels in the NHS.
- Second, consistent failure to meet targets brought considerable external attention and reputational damage.
- Third, the hierarchy of urgency associated with the targets provides an insight into the varied ‘buffer capacity’ of this boundary. The RTT of 18 weeks had a longer time scale and therefore some buffer capacity could be created in comparison to the 4 hour ED target. When a hospital operated very close to the target, such as 97.2% of the expected 98% ED target, there was no ‘buffer capacity’ related to that boundary.
- Fourth, due to the location of the marginal zone boundary close to the failure

boundary, considerable effort was expended to keep the OP away from the target failure boundary. Some of the efforts have been described as ‘gaming’ (Bevan and Hood, 2006). The compensating actions required to avoid target breaches, such as generating ‘buffer capacity’ by opening additional beds and outlying medical patients on non-medical wards, created a number of interactions with the other boundary constraints. These interactions created dynamics that will be explored in more detail in the next two chapters.

The examples of the dimensions derived from the data are summarised in Table 6.7:

Construct	Dimensions	Examples from the data
Target failure boundary	Visibility	Nationally set, performance managed at national, regional, hospital and department level.
Target failure boundary	Movement	Almost no movement of marginal zone boundary.
Target failure boundary	Location	Marginal zone boundary location was close to the failure boundary.
Target failure boundary	Marginal zone (buffer capacity)	Limited buffer capacity for RTT, often none for 4 hour target.

Table 6.7: Dimensions of the SWE ‘Financial failure boundary’ construct with examples from the data

In terms of the SWE model, it is possible to depict the ‘target failure boundary’ as being a dominant factor in the decision making and actions within the hospital. This dominance can be explained because the target boundary was arguably the most ‘visible’, clearly ‘located’ with least movement of the marginal zone boundary and had almost no buffer capacity.

6.7 Summary

Rasmussen (1997) developed the interacting three boundary SWE to depict the socio-technical context of safety. The boundaries of the envelope define the constraints within which a system operates. Rasmussen (1997) does not go into detail about the dimensions of the boundaries, nor how different levels in an organisation might perceive them. When the model is contextualised and applied in the case studies, the

dimensions of each boundary were perceived in slightly different ways, which generated different responses.

For example, the unacceptable workload boundary had some flexibility and was less clearly defined further away from the 'sharp end' (Cook and Woods, 1994, Flin et al., 2008). The patient safety failure boundary was only partially defined and measured. Breaches of that boundary were not always noticed. When staff normalised to certain situations, such as patient falls, the marginal zone patient safety boundary was moved outwards to where a new safety 'norm' is established (Vaughan, 1996). The financial failure boundary had some buffer capacity given that success was judged over the full year, although monthly monitoring occurred. Front line staff were less concerned about finance than about workload and patient safety. The target failure boundary, which is contextually specific to the NHS, was shown to be the most 'visible' boundary with little or no buffer capacity. The lack of flexibility associated with the target failure boundary created certain behaviours within the system. The timescale and therefore the hierarchy of urgency, within which human agents took actions and made choices to avoid target breaches, appears to have been much shorter than in relation to the other boundaries.

The dimensions of the different boundaries are illustrated in Figure 6.9

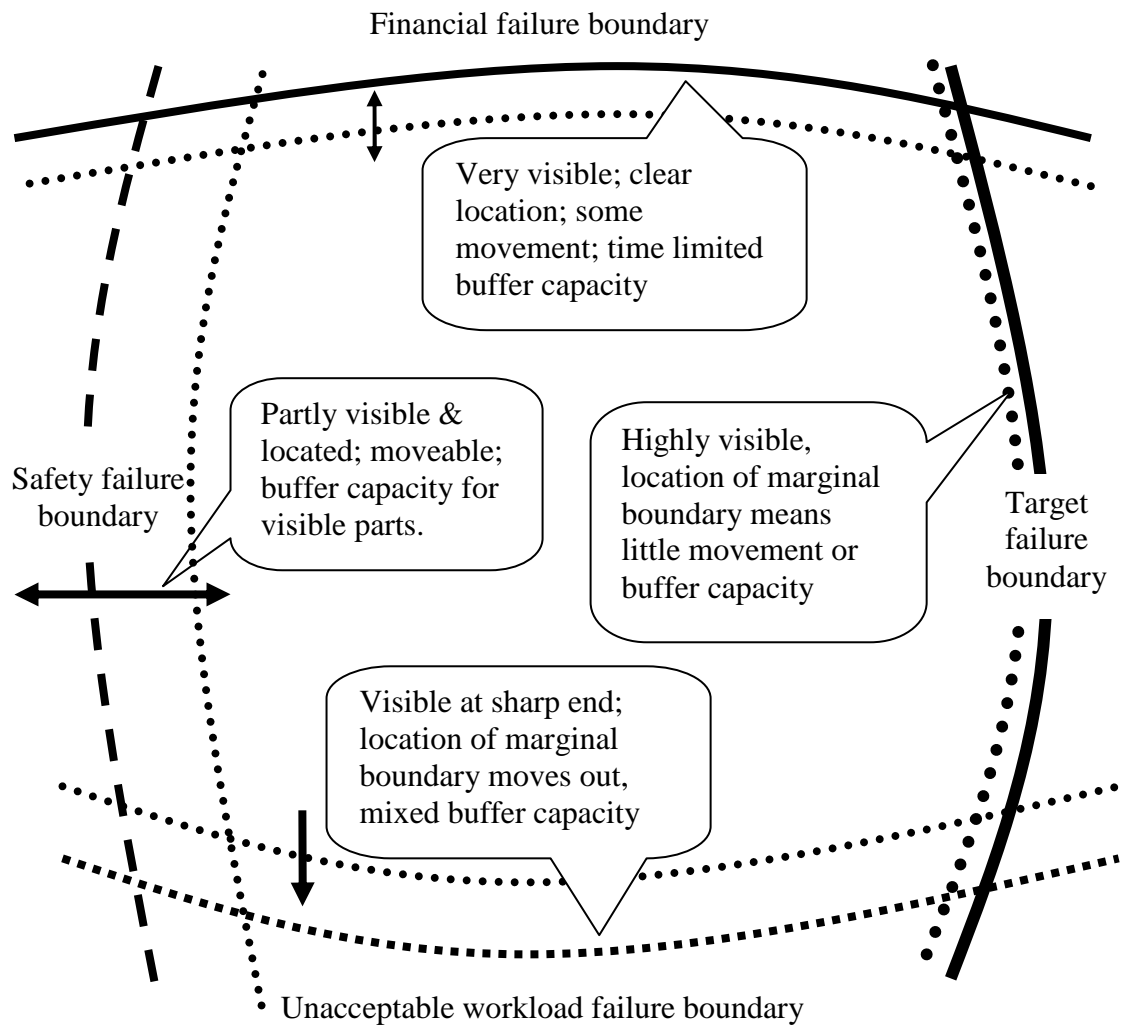


Figure 6.9: The SWE with boundaries dimensions

The boundaries depict the constraints within which the system is designed to operate (Woods, 2006). The boundaries create competing pressures on the OP of the system. What is shown by the case study data is that at periods of high demand, the dimensions associated with a boundary determined, in part, the priority given to actions to keep the OP from breaching the boundary. Interestingly, the more visible and less movable the marginal boundary, the greater the priority was given to the compensating actions that needed to be taken.

The boundaries of the SWE are only one part of the model. The pressure created by the boundaries is linked to the gradients construct of the model that sets the conditions within which the OP operates. Part of what has been considered above is related to the gradients. For example, the Department of Health's Operating Framework documents can be regarded as contributing to the gradients related to all four boundaries that

influence the movement of the OP. In the next chapter the dynamic interactions created by the gradients influence on the OP are explored.

Chapter 7 – Investigating the Gradients and Operating Point

7.1 Introduction

This chapter explores the dimensions of the ‘gradient’ and the ‘operating point’ constructs of the extended SWE (v3). The gradients depict the pressures on the OP. The issues explored in this chapter are both the pressures and the system behaviour as depicted by the OP (Figure 7.1).

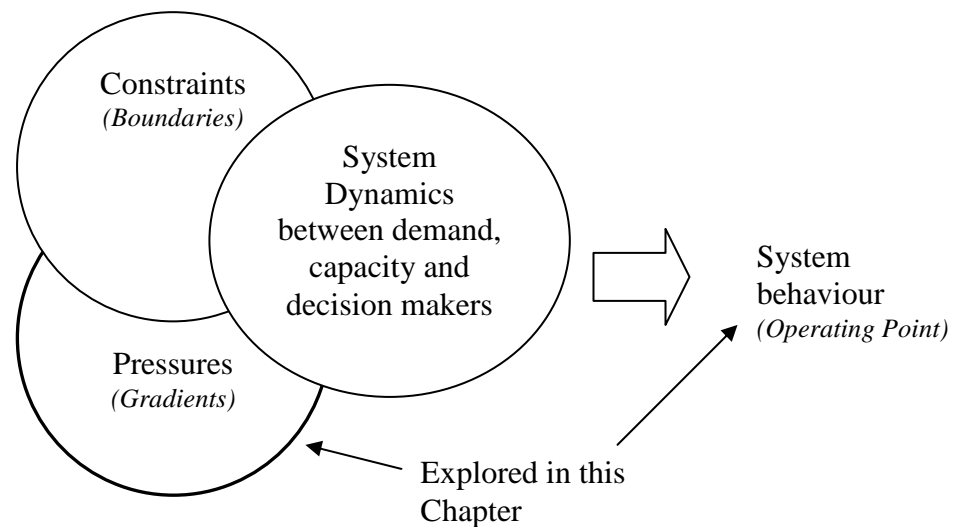


Figure 7.1: Combination and interaction of construct sets depicted by the SWE model

The dimensions of the gradient and OP constructs are derived from the literature in Chapter 3 and for ease of reference are presented in Table 7.1. The exploration of the dimensions is conducted through the analysis of three events from CS 1 and 2. The first event from CS 1 was an outbreak of a sickness virus in the hospital which closed a number of wards for just over a week. Conceptually this event depicts a sudden perturbation when the OP breached the patient safety boundary, which generated actions and interrelationships between the competing gradients. The second event from CS 1 was a surge in emergency demand that lasted more than two weeks, which created significant operational problems. This event illustrates ‘continuous stress’ on the hospital, the shift of the marginal zone boundary location, and the ‘drift’ of the OP towards the patient safety failure boundary. The third event from CS 2 was the flow of emergency patients through the ED and MAU on one particular day. The data provides

an insight into how the dynamics of the gradients and OP movement, at a hospital system level, can impact at the micro patient experience.

The three events provide data on the ‘movement’ and ‘location’ dimensions of the OP construct of the SWE (v3). As noted in Chapter 3, when small movements of the OP is combined with the location remaining inside the envelope, then the system is conceptualised as being highly reliable (Cook and Rasmussen, 2005). Alternatively, when the movement of the OP is large and the location is in the marginal zone, or breaching a failure boundary, then it is regarded as a low reliability system. The reader is reminded that a resilient system is conceptualised as one where the OP can remain within the boundaries of the SWE in the face of perturbation or continuous stress.

The gradients linked to the boundaries apply ‘pressure’ on the OP. Rasmussen (1997) argues that the gradients drive the OP towards the safety failure boundary. As such the gradients provide a ‘downward pressure’ on the OP, moving it away from the boundary associated with that gradient. However, in Chapter 3 and 5 the gradient construct is developed to take account of the ‘upward pressure’ created by the responses of the agents to the multiple interactions associated with the different gradients. The response by agents to the downward pressure is described as ‘cross scale interactions’ (Woods, 2006). Therefore, it is helpful to explore both the ‘pressure’ and the ‘scale’; namely establishing whether the pressure is downward or upward. By examining the three events, understanding is gained into the ‘scale’, and ‘pressure’ dimensions of the ‘gradient’ construct. It is argued that the constructs of the model provide insights into the characteristics in NHS hospitals in England, and how they potentially influence patient safety.

Construct	Dimension	Quotes from the literature
Operating Point	Location	'...its distance from the marginal boundary' (Cook and Rasmussen, 2005)
	Movement	'At the moment there is little or no work directed at characterizing the location or movement of the OP of the system' (Cook and Rasmussen, 2005)
Gradients	Scale	downward and upward scale interactions (Woods, 2006)
	Pressure	'continuous pressure' of the safety counter gradient 'compensating the functional pressure of the work environment.' (Rasmussen, 1997)

Table 7.1: Dimensions of the SWE 'operating point' and 'gradients' constructs derived from literature

The events are described and the evidence from the data collected is presented about the response by staff in the hospitals. The data presented has been reduced from interview transcripts, field notes of observations and hospital administrative data. Data source codes are shown in square brackets [].

7.2 Event one – Norovirus

In CS 1, an airborne sickness virus (Norovirus) in one ward spread to other wards. Within a few days the virus was present amongst patients and staff on twelve out of forty nine wards [CS 1: Doc 4]. The Control of Infection procedures for such an outbreak state is that the affected wards or bays are closed to new admissions. Only patients going directly home are allowed to be discharged. Therefore, the transfer of patients to community hospitals or nursing homes was stopped, which slowed the discharge rate. The reduced discharge rate and the closed wards, created a perturbation in the normal running of the hospital. The virus closed eight wards to new admissions. Therefore, there was reduced bed capacity to accommodate both the emergency and elective patient admissions. Visitors were not allowed on the infected wards and staff movements were restricted to essential visits only.

In conceptual terms the sickness virus created a rapid movement of the OP. It is argued that due to the level of harm to patients created by the virus, the OP breached the patient safety boundary (Figure 7.2). The downward pressure from the gradient towards production associated with the target failure boundary created the conditions where the

rapid spread of the sickness virus was more likely. As noted, the Annual Control of Infection Report [CS 1: Doc 5] implicated the high bed occupancy and movement of patients as contributing factors to the spread of Norovirus. The high bed occupancy, which created large number of outliers and the patient movement, was the result of pressure from the gradients towards efficiency and production. Further evidence of this is presented below.

The buffer capacity of empty beds able to be deployed to meet the initial Norovirus outbreak was minimal due to the high bed occupancy and number of medical outliers. Conceptually, there was a rapid decompensation as the buffer capacity was quickly exhausted (Miller and Xiao, 2007), resulting in the OP breaching the patient safety failure boundary.

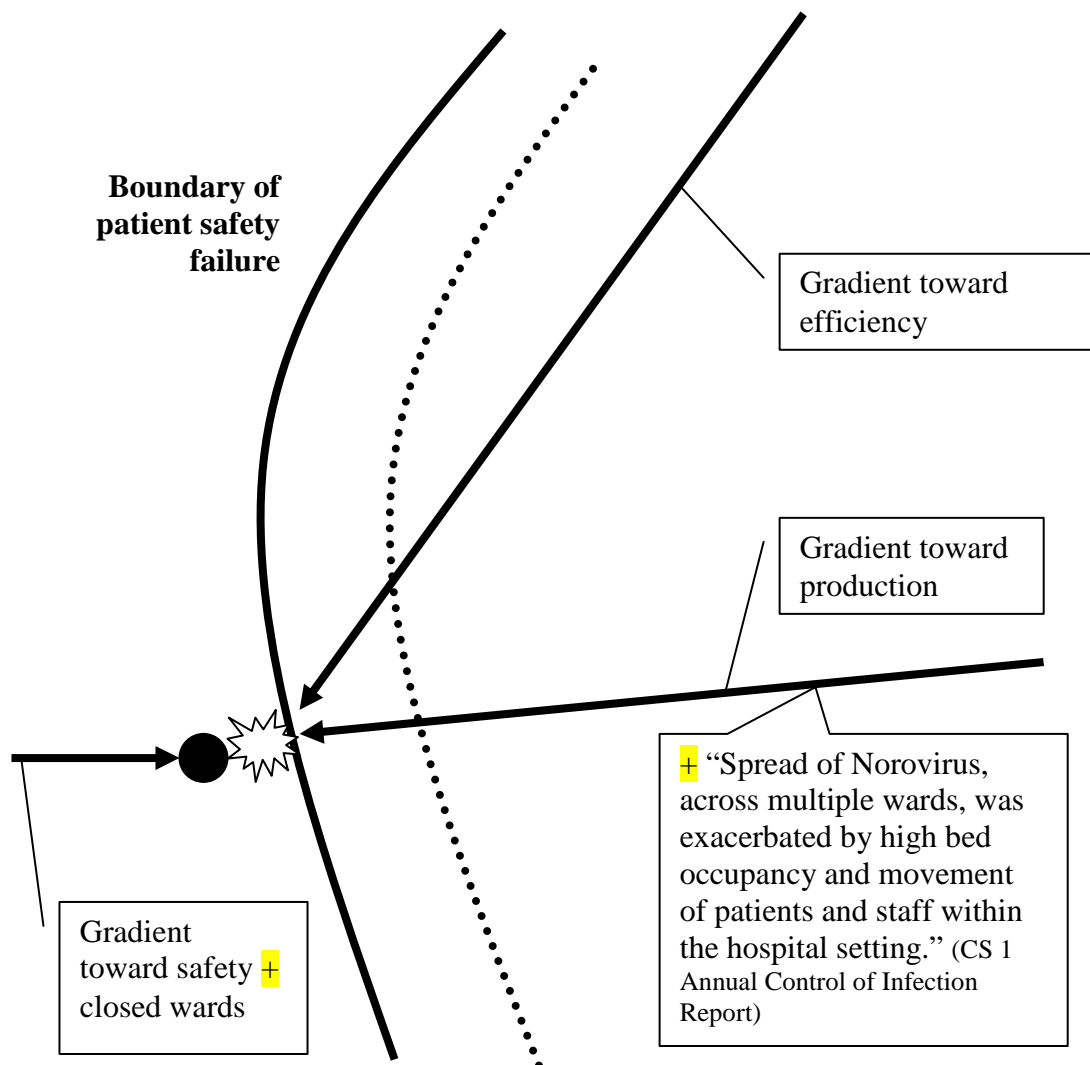


Figure 7.2: The OP breaching the patient safety failure boundary due to the rapid spread of the Norovirus in CS 1
 (+ = downward pressure moving the OP away from associated boundary)

The compensating action of closing wards to contain and eradicate the virus can be conceptualised as the patient safety gradient pushing the OP back inside the envelope. An analysis of the response to the outbreak is set out in the next section and the insights into the dimensions are identified.

7.2.1 Response

The on call Director took the view that it was not possible to close the hospital to new admissions for two reasons. First, patients would still arrive and staff would find it hard to turn them away and second, neighbouring hospitals did not have sufficient capacity if emergency patients were diverted to them [CS1.15]. A major focus of attention was the need to contain and remove the virus. Ward closures were strictly adhered to.

“...the first thing that we have done is very clearly not compromised the areas that are infected so those areas have got to be closed and they are quite clearly closed, there’s no transfers out from them, there’s no admissions into them so the patients can be discharged from them but we are not moving them anywhere else so they have been very much locked down and for us this week that’s been kind of 8 ward’s worth.” (Manager CS 1) [CS1:RP; DM; Capacity]

Over the following week a thorough cleaning programme of areas was used prior to reopening the wards. The hospital was closed to all visitors until the wards were reopened. A press release was given out to explain the situation [CS 1: Doc 1.4].

At the same time the reduced bed capacity meant that there was a danger of the OP breaching the target failure boundary. The ‘pressure’ from the production gradient related to the target failure boundary then influenced decision makers to open beds in a reactive manner and spread medical and nursing staff more thinly to cover those areas. The dynamic interactions create the situation where the OP moves over time as illustrated by Figure 7.3.

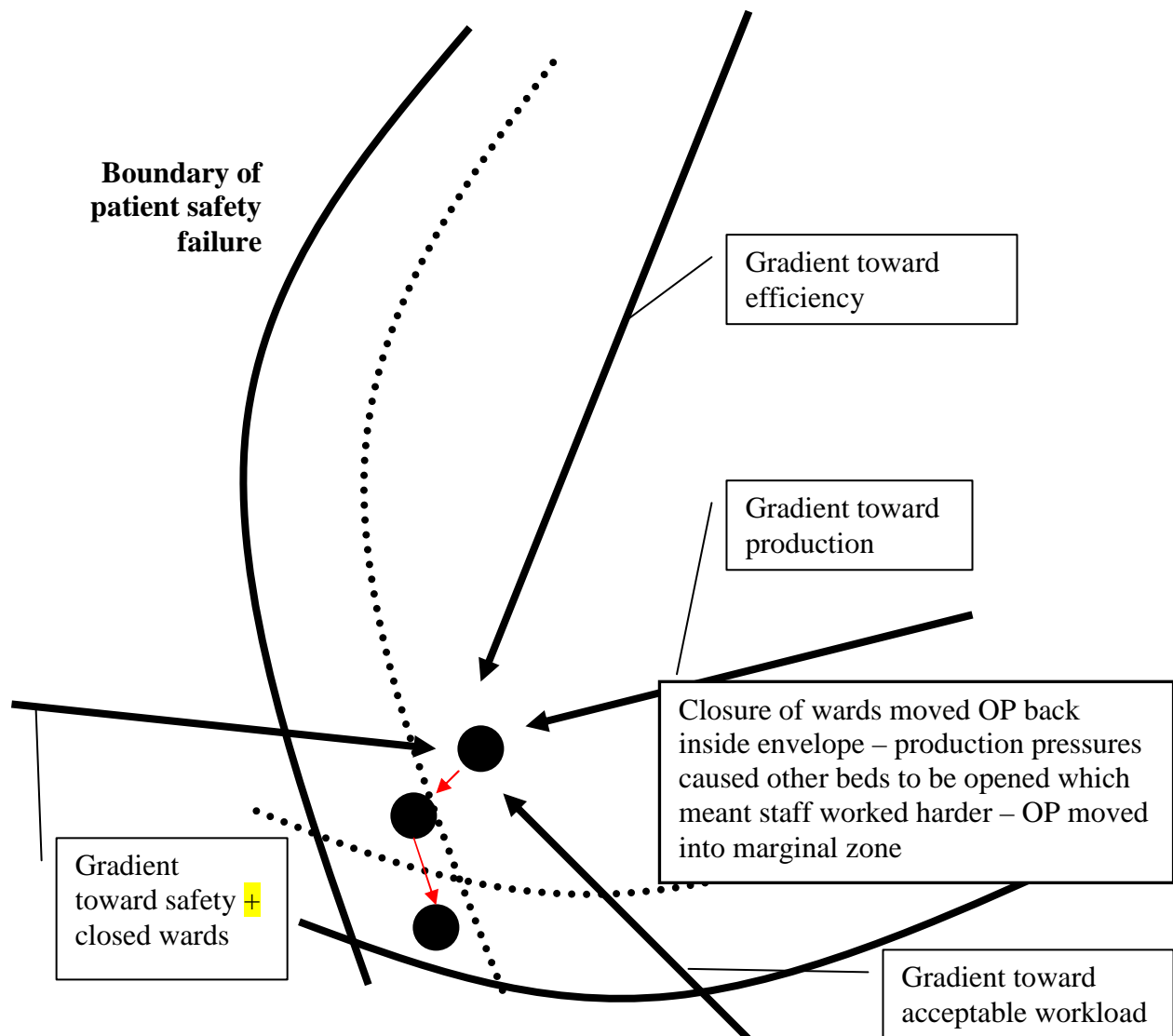


Figure 7.3: Movement of the OP following the outbreak of the Norovirus

The next section provides evidence of the actions taken in response to the pressures exerted by the different gradient pushing the OP away from breaching the failure boundaries.

7.2.1.1 Opening additional inpatient beds

A number of compensating actions were taken within the hospital to generate new staffed bed capacity. These beds were needed to maintain the rate of admissions and avoid the OP breaching the target failure boundary. The actions included opening an area normally used as a pre-admission area in orthopaedics (normally only open during the working day). This small 13 bed area was situated at the other end of the hospital from the medical wards. It was not equipped as an inpatient area; there was no drug trolley or inpatient beds [CS 1.15]. Patients already in medical beds, identified as being

well enough, were then moved to this area. A further area of 6 beds, known as the Clinical Decision Unit (CDU), was brought into use in an area beside the ED. This was a prefabricated building with no natural light. This was used for medical patients. Areas within medical and surgical wards that were normally used as day case beds were brought into service as inpatient beds.

“So we’ve looked at every kind of area of capacity within the organisation and been a bit creative about how we can keep work going but by doing it in different ways.” (Manager CS 1) [CS1: RP Capacity]

7.2.1.2 Increasing the rate of medical outliers

The OP can be conceptualised as being in the middle of the envelope if all the gradients apply equal pressure. In such a situation all the patients would be in the correct bed within the target time, there would be sufficient staff on duty and the budget would be in balance. When there are a number of medical outliers the assumption is made that the OP is being pushed towards the patient safety failure boundary.

A daily requirement for the midday bed meeting was for nursing staff on the medical wards to identify those patients they consider suitable to be moved to a surgical ward.

“I mean you know if I’m honest we have been asked to number our patients 1 to 28 with 1 being the most suitable to outlie and 28 being the least suitable and if necessary they would work down that list.” (Nurse CS 1) [CS1: AD; Medical Outliers]

By moving medical patients into surgical beds it provided the capacity within the medical wards to accept transfers from the Emergency Medical Unit (EMU). In turn then EMU could accept admissions from the ED. To achieve the 4 hour target in ED there has to be a constant supply of beds being found in both EMU and the medical wards to be able to admit patients. Staff planning to move medical patients onto surgical wards usually identified those patients who are fit enough to be sent home within the next day or two. However, at times of high demand or where wards are closed, this was not always possible and less suitable patients are moved. This was explained by one of the site managers:

“We would normally be looking for patients who are going home the next day or within 48 hours that have got a definite plan. In times of extreme pressure we would be looking at patients that could be, but don’t have a definite discharge plan but could be safely nursed in another area.

And who makes that judgement?

Well we ask that the nurses on the wards the problem is when we had the sickness virus the number of wards that we could actually outlie from was

diminished so we were targeting those areas more and more daily. Now the nurse in charge of that ward area was asked to put them in priority order you know on one particular day I recall that one particular ward was asked to give us 12 names and prioritise those names as to you know

And how many of those ended up moving?

The whole lot.” (Site Manager CS 1) [CS1: AD; Medical Outliers]

As can be seen from this quote there is a sense of normalisation (“We would normally...”) to having to find patients to move. Such actions have a potentially detrimental effect on patients. However, when there was a lack of ‘buffer capacity’ to meet the various targets as identified in Chapter 6, then the ‘pressure’ from the target boundary gradient meant that staff took actions that push the OP away from breaching the target failure boundary. There were potential implications for patient care from those actions, which can be viewed as pushing the OP towards the patient safety failure boundary.

7.2.1.3 Implications for the patients

The doctors were not routinely part of the decision making process to outlie patients even though one of the consequences can be that the care of that patient is handed over to another medical team {CS 1.17}. Staff who received medical outliers do not always have the expertise to look after them as pointed out by a specialist doctor:

“Because often after the admission you recognise that somebody is a specialist patient, they have got ischemic heart disease and need an angiogram and they won’t get listed for the angiogram until they are on a cardiology ward. They might be on a waiting list but they will be delayed if they are not on a cardiology ward. If they are a respiratory patient they will not get the specialist consulting review. If they are a diabetes patient and they are diabetic ...and they have gone to a non-diabetic ward they may be there for you know they will be there longer because they won’t have their insulin changed or they might not have their insulin changed appropriately and also somebody with a diabetic foot somebody might not appreciate the seriousness of that because they are not trained for that disease.” (Doctor CS1) [CS 1: AD; Medical Outliers]

Staff consistently expressed concerns about the care received by medical outliers. From the interview and observation data an inference is made that outlied patients are at higher risk of delays in treatment with potential associated effects on their safety. It is apparent that at times of high pressure patients could be moved at all hours during the night. The following is an extract from the Field Notes having copied the ED ‘Breach’ reports from the Site Managers working during the night: (ward names have been anonymised)

ED Breach of Four Hour Target Reports from overnight: 5/6 January 2009

“00.53hrs – moved patient from Ward H to Ward O. Moved EMU side room to Ward H side room. EMU side room needed cleaning. When S/R ready on EMU, ED explained that patient needed 1:1 nursing. EMU could not give 1:1 nursing; patient went to ITU”

“0216hrs had to transfer patient to side room on Ward L from EMU. EMU side room had to be cleaned before patient could be admitted.”

“0256hrs Patient vomiting. Unexplained cause. Needed side room on EMU. Had to move patient out of side room on Ward T. Transfer patient from EMU. Clean EMU side room before patient could be admitted.”

[FN 4.62-64]

When patients were potentially or known to be infectious, side rooms were required to isolate them and prevent the spread of infection. The control of infection was regarded as being the priority patient safety issue. Side rooms were a scarce resource. The demand for them required a process of constant reassessment of priorities and reallocation to higher risk patients when appropriate, even if that is in the middle of the night.

“I think it’s just always wrong to have outliers always wrong to move them unnecessarily you know we have had some truly catastrophic cases as you are probably aware of people in their nineties perhaps moved four or five times and you really you look back at the case and you think how did that happen you know at some point somebody would have said enough is enough but it’s because you’ve got a group of bed managers doing their best with a number of really bad conflicting alternatives making people stack up in ED when perhaps they are extremely vulnerable or on EMU or, and try to meet the infection control needs because that’s the other dimension you know when you, it takes a lot of flexibility out of the system if you have to move people into side rooms and so on and perhaps you know certain bays you can’t move people out of and so on.” (Doctor CS 1) [CS 1:AD; Medical Outliers]

There was wide agreement among staff that outlying patients was far from ideal. It is inferred that the ‘downward scale pressure’ on the OP from the target and financial boundary gradients left them little if any choice. Therefore, moving large numbers of patients to wards not specialising in their care appeared to have become a ‘normal’ means of creating ‘buffer capacity’ for the target failure boundary. The patient safety consequences did not appear to be as important as the necessity to be able to admit patients in the timescales required. Conceptually, the gradient exerting pressure related to the control of infection was acted upon by closing wards. At the same time, the patient safety concerns of outlying patients were outweighed by the pressure exerted by the gradient towards greater production associated with the target failure boundary.

In CS 1, one of the triggers to initiate the ‘escalation plan’ was having more than twenty medical outliers [CS 1: Doc 3]. The escalation plan brought more senior decision makers into the bed meetings in an attempt to keep the hospital running smoothly. Therefore, conceptually it can be argued that, when there were more than twenty outliers it indicated that the OP was inside the marginal zone of the patient safety failure boundary. Figure 8.13 suggests that in CS 1 it had become ‘normal’ for there to be more than twenty medical outliers. As suggested by Wheeler (Wheeler, 2003), a moderate but sustained effect on the number of medical outliers was observed during the sickness virus period with ten consecutive data points above the mean but within the upper control limit (Figure 7.4).

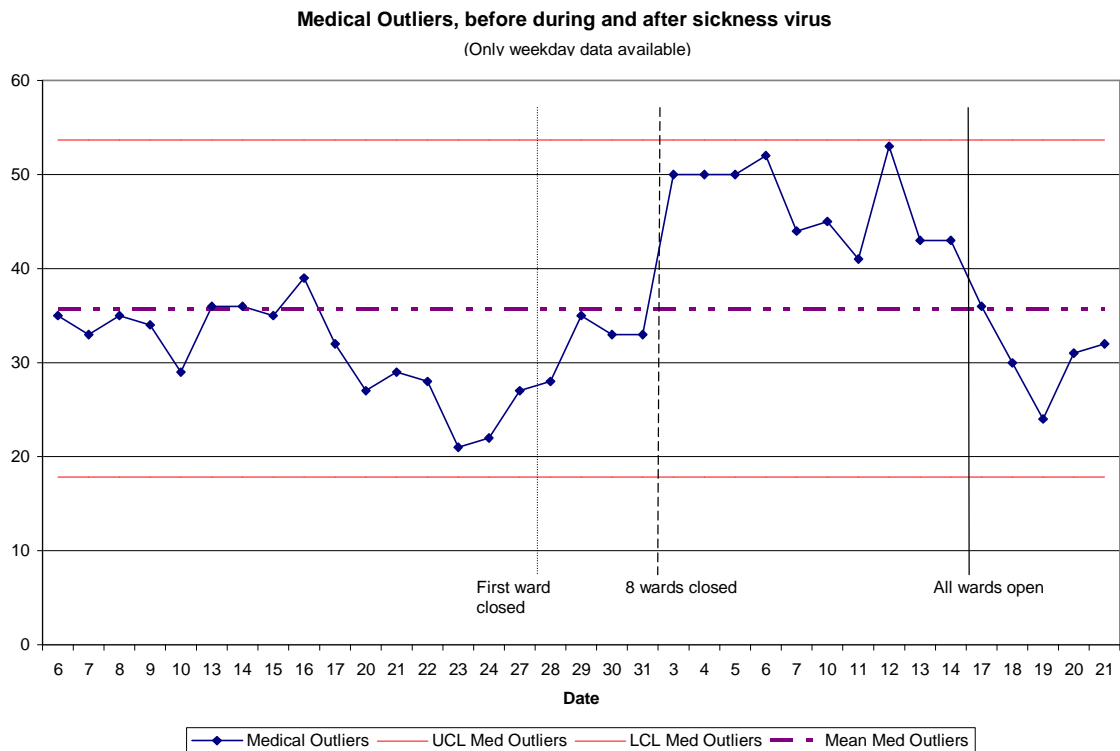


Figure 7.4: Statistical Process Control Chart - Medical outliers before, during and after sickness virus

At the bed meetings staff were under pressure to find bed capacity by encouraging discharges. Site managers asked staff to identify patients approved by the doctors for discharge so they could “sit them out”. This meant moving the patient to a chair to free that bed for the next patient. Doctors were encouraged to conduct extra ward rounds to see if more patients could be discharged. Conceptually, these actions are in response to the ‘downward pressure’ of the gradient towards increased production to meet the target requirements. The need to keep the ED patients being admitted within the four hour

target is explored next as an illustration of the ‘pressure’ to keep the ‘location’ of the OP away from the target failure boundary.

7.2.1.4 Keeping ED flowing

The managers and site management team worked hard with ward based staff to juggle the competing demands. During the Norovirus there was no question of not achieving the waiting time targets for patients [FN 1.32]. Often staff justified the four hour waiting time target in ED as being good for patients:

“I don’t disagree with the like for example the ED target I mean it’s good that we do have a four hour turn round because you don’t want patients laying on a stretcher for more than four hours because that’s not in, in the extreme you are doing patients harm you know that’s, so it’s important to keep the flow going.” (Nurse CS 1) [CS 1: AD; Flow]

However, the consequences of ensuring that the four target was achieved placed pressure on other parts of the system and particularly on EMU as explained by one of the nurses working there:

“...overwhelmingly the momentum is driven by demand and capacity and flow and you know you learn that concept very quickly when you work on EMU. We have to manage the admission process through the EMU we take those from the GP and also the emergency department. At this time of year we are moving into the 55/60 every 24 hours, that’s a large number coming through, we are a 31 bedded unit.” (EMU Nurse CS 1) [CS 1: AD Flow]

Each morning EMU sought to discharge some of the overnight admissions to provide beds for emergency patients sent in by GPs. However, if demand was greater than the number of empty beds the ‘medical take’ (emergency admissions from GPs) was diverted to the ED. This became a daily occurrence during the working week of CS1.

“Normally by about midday, it depends on the number of calls, but normally by midday if we have only been running with one bed, you know one patient in and one patient out normally by about midday/ 1o’clock ED are having to take because we are waiting for the GPs you see once they start their clinics you get a couple, probably one or two up until midday may be one or two from ED not much but as soon as GPs get off house calls that’s it then, once they are out on their house calls their numbers start shooting up and then you start to take quite a large number through. So roundabout, they start their house calls about by midday and complete them by about 3 o’clock so between 12, 3 and 4 o’clock you start taking a large quantity of GP referrals.” (EMU Nurse CS 1) [CS 1AD; Emerg Adms]

Observation data, confirmed in interviews, shows that at times during the afternoon ambulance staff queued with their patients in the ED corridor because the department

was full [CS1.28; FN 2.38; 56]. Pressure was then on EMU to discharge or transfer patients in order to receive patients in a timely manner from the ED. The movement of patients out of EMU allowed ED to accept patients from the ambulance crews. The drive to admit patient from ED within the four hours did have other implications, as one doctor observed:

“the four hour wait spews people out at any stage of investigation off into any bed that’s available at the time.” (Doctor CS 1) [CS 1:AD; Emerg Adms]

The implications for the doctors of patients being moved rapidly to beds scattered across the hospital was that they could lose track of patients causing delays in treatment. The medical team admitting patients relied on a hand written sheet of patient names and locations kept up to date by the on-call Registrar or Senior House Officer [FN 3.24; CS 1.31; CS 2.15]]. They struggled to keep the sheet updated when patients moved to unexpected locations.

7.2.1.5 Not cancelling elective patients

One of the actions that could have been taken to reduce the ‘downward pressure’ from the target boundary gradient on the system was to cancel elective admissions. Most elective patients occupy surgical beds. All elective patients have a maximum waiting time (RTT target). Achievement of such targets was an important issue for the hospital. When asked about the importance of the RTT target, one manager responded:

“Well it’s yeah it’s one of the kind of key targets that as an organisation we are looking to achieve and we have kind of signed up to it, a 13 week referral to treatment time by the end of well by, through March of 2009 and so we are on the kind of trajectory to do that. I think you know relatively as an organisation 18 weeks and increasingly 15 weeks isn’t a big issue for us, 13 weeks is within orthopaedics all of those targets were a big issue. So you know the pressure to kind of keep things going within all of those areas is really quite considerable. I think within some of the surgical specialties we’ve got some flexibility about looking at deferring admission but it’s clearly a kind of key priority that you know and a cancellation is kind of you know a little way down the line. ... We are still managing to put the work through but it’s being done by some of the creative solutions that are kind of suggested around...” (Manager CS1) [CS 1:SWE; Targets]

From a conceptual view this quote emphasises the importance attached to the gradient associated with the target failure boundary. It can be inferred that there was some ‘buffer’ in certain specialities that were ahead in meeting the 18 week target. That capacity could have been used to cancel patients and relieve some of the pressure on the

staff and beds during the sickness virus. However, there was also a strongly held view that such elective patients should not be disadvantaged by being cancelled, as their treatment needs were of equal importance to the emergency patients [FN 1.32; CS 1; AD; Elect Adms]. Analysis shows that during the period of the sickness virus the number of elective admissions was slightly higher than for the same period the year before (Figure 7.5).

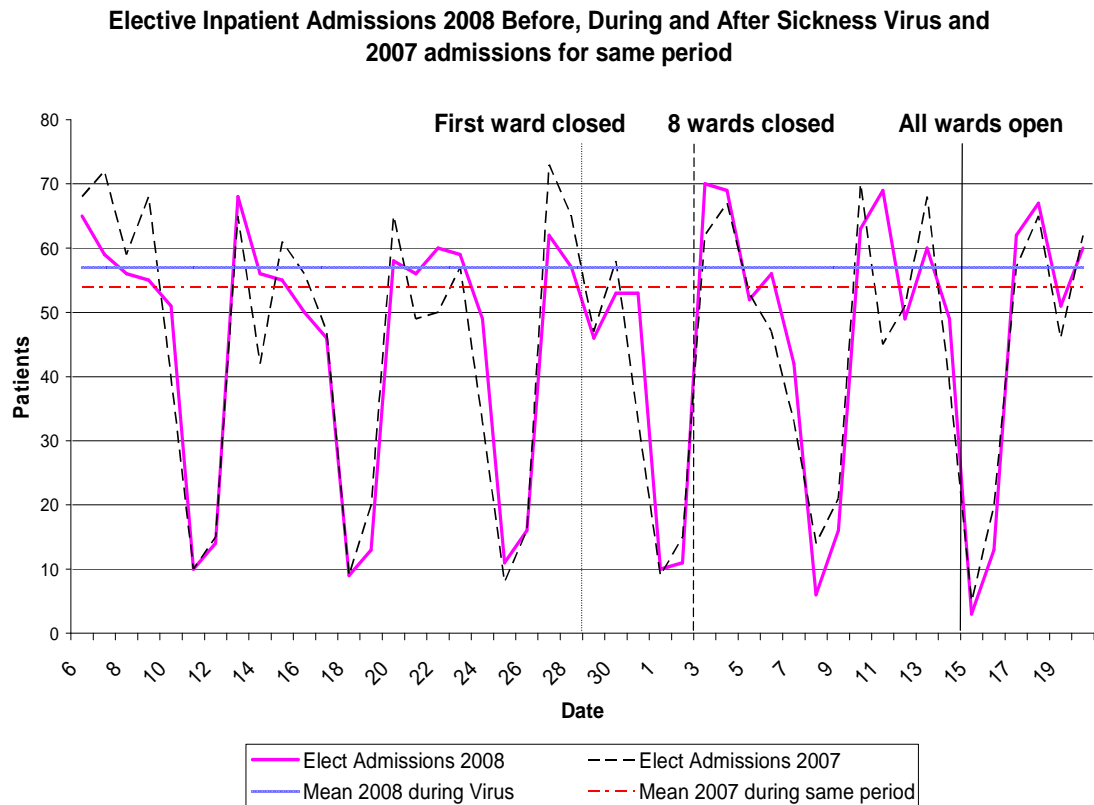


Figure 7.5: Elective Admissions 2008 Before, During and After Sickness Virus and 2007 over same period

In conceptual terms, the managers choose *not* to exercise ‘upward pressure’ on the gradient towards greater production by cancelling elective patients. The consequences of continuing to admit all the elective cases meant that there were more patients than beds available at certain points in the day.

“So the throughput is phenomenal yet they haven’t got any beds for these patients and yet they are having to start theatre lists without having you know a bed for a patient. So it’s each day trying to work how we can best make, it’s safe for the patients to go to theatre so the staff feel that they are looking after a full ward plus also a corridor of patients.” (Surgical Nurse CS 1) [CS 1: AD; Elect Adms]

The situation was made worse by the patients to be operated on that day often being asked to arrive at 8am. This created a peak in demand for beds at a time when the patients to be discharged that day had not been organised. Patients were observed waiting in corridors, waiting rooms on wards and being prepared for theatres in offices [FN 2.40-50]

Part of the juggling act for the site management team was finding beds for surgical patients post operatively, whilst at the same time finding beds for emergency admissions [FN 1.44; 2.64]. The Recovery Room in theatre was often used as a holding area for patients waiting for a bed after their surgical procedure. On one occasion it was observed that a Recovery Nurse was asked, by a ward without any empty bed, to arrange the discharge of a patient home directly from the Recovery Room [FN 2.52]. She was unsure how to do this and required assistance with the process.

Conceptually, the downward pressure from the gradient to keep the OP away from the target failure boundary can be inferred to be high. The consequent actions taken to meet the waiting time requirements can be regarded as ‘moving’ the OP into the marginal zone if not through the patient safety failure boundary. As noted in Chapter 6, harm to patients is not always obvious. Decision makers can therefore ‘stretch’ the system believing that patient safety is being maintained. When the hospital was overcrowded there was also an impact on the workload of staff which is explored next.

7.2.1.6 Staff workload

The daily bed meeting often included a subsequent meeting of nurses who kept each other informed about potential and actual staffing difficulties [FN 2.13]. Risk assessments were undertaken on staffing levels at various times during the day. Professional judgment was used to decide the allocation and mix of trained (registered nurses) and unregistered or agency staff to the wards. This resulted in staff being allocated to wards that they are not familiar with. During the sickness virus outbreak staff were classified as either ‘dirty’; having worked on a closed ward, or ‘clean’ and redeployed accordingly. Nursing and medical staff were moved within the hospital to cover the newly opened bed capacity and temporary staff were sought from the Nurse Bank (internal provider of flexible staff) and Agencies (private sector providers of staff at higher cost) [FN 2.34].

Considerable effort was expended several times each day to manage the staffing levels to keep the bed capacity operational. No areas were closed due to lack of staffing. Some areas were flagged as being at the limit of the staff's ability to cope. In certain instances newly admitted acutely ill patients were 'deflected' from their intended destination and admitted to other areas to protect stretched staff. Patients were sometimes admitted or transferred to areas where the staff were not specialist in their treatment. Medical staff along with nurses fell victim to the sickness virus. Doctors who worked on the wards affected by the virus were not allowed to visit 'clean' wards. The medical staffing was therefore disrupted. The doctors in the medical specialities were deployed with often just one junior doctor per ward rather than the normal two [CS 1: SWE; Staff workload].

Therefore, during this time of disruption the staff workload increased. Day case areas were used for inpatients but the patients requiring the short stay procedures were not cancelled. The consequence was that nurses had a much higher workload:

“...it's a very fast turnover here and the staff are having to clear beds, get people sat in waiting rooms so they have effectively if you've got like a double set of patients because they've got the ones that we are then putting into that bed, the ones that they are still looking after who are in the waiting room waiting for discharge that they are constantly checking on, doing their observations, making sure they are safe, so that you know they are safe. So you've got you know almost a double compliment of patients all the time.” (Ward Sister CS 1) [CS 1: SWE; Staff workload]

Senior nurses in particular focused a lot of effort on keeping staffing workloads at what they regarded as a safe level. Their actions can be conceptualised as applying 'downward scale' pressure on the gradient to keep the OP away from both the workload and patient safety failure boundaries. However, the reality of the situation, especially for the doctors and nurses who had to move wards, was that they worked harder and sometimes in strange environments, which inadvertently, makes them more likely to make an error (Reason, 1990). The consequences for individual staff can be conceptualised as 'upward scale' pressure on the gradients 'moving' the OP 'location' towards the intersection of the unacceptable workload and patient safety boundaries.

Finance

The strong financial position of the CS 1 hospital meant that additional resources were spent to pay for staff overtime, bank nurses and agency staff. Budget holders were

expected to account for their budgets and the reasons behind additional expenditure at the routine performance management meetings. It is clear from interviews with senior managers that financial considerations were not to compromise the safety of patients in terms of staffing the hospital during periods of disruption [CS 1; SWE; Finance].

Conceptually, the OP was 'located' well away from the financial failure boundary and during this time of perturbation the 'downward scale' pressure on the gradient towards efficiency was relaxed to allow additional expenditure.

7.2.2 The Operating Point and Gradient dimensions

The description of actions, views and consequences can be related conceptually to the SWE (v3). It can be argued that the OP breached the patient safety boundary due to the high number of infected patients and the number of medical outliers. The actions to resolve the rate of infection brought the OP back inside the envelope. However, despite relaxing the financial gradient, the staff workload was higher than normal for nurses and doctors due to high sickness levels. The OP therefore can be regarded as being 'located' in the marginal zone at the intersection between the unacceptable workload failure boundaries. The pressure from the target failure boundary gradient had the consequence of making staff take certain action that increased their workload. More beds were opened, elective patients were not cancelled but allowed to be kept waiting in corridors, the number of medical outliers increased, staff were redeployed and doctors encouraged to find more discharges. Therefore, the OP 'location' was kept away from the target failure boundary.

The dimensions of the OP are the 'movement' and 'location' (Table 7.1). When the OP was located inside a marginal zone then there is a danger of a resulting breach of the corresponding failure boundary. The sickness virus event described above illustrates the OP breaching a boundary and then being pushed back inside the envelope. The OP is then subject to considerable pressure to keep it away from the target failure boundary. The size of the 'movement' of the OP can provide an insight into the stability of a system (Cook and Rasmussen, 2005). When there are large and sudden movements of the OP the system can be conceptualised as unstable. At the outbreak of the sickness virus event it is argued that there is a sudden movement of the OP outwards through the patient safety boundary. The actions taken to isolate the infection moved the OP back inside the envelope. However, the pressure from the target boundary gradient moved it

back into the marginal zone at the intersection for both the patient safety and unacceptable workload boundaries.

From the analysis above it has been noted that there was an 'upward' and 'downward' pressure. For example, senior nurses applied 'downward' influence (pressure to move the OP away from workload failure boundary) by bringing in extra staff and moving nurses around. However, the actual work experience of many staff applied an 'upward' influence (moving the OP closer to the workload failure boundary) due to them overworking. During this event the gradient towards efficiency from the financial failure boundary exerted low 'pressure' in comparison to the apparent high 'pressure' given to the gradient related to the target failure boundary.

Conceptually, there was a degree of 'normalisation' to the virus outbreak and the consequence actions and ensuing conditions. This was exemplified when a report about the Norovirus outbreak was brought to the Governance Committee of the hospital some weeks later. Despite the considerable disruption observed, there were no comments made about the report and the meeting moved on to next business [FN 1.74].

The data from this event is conceptualised and displayed in Figure 7.6.

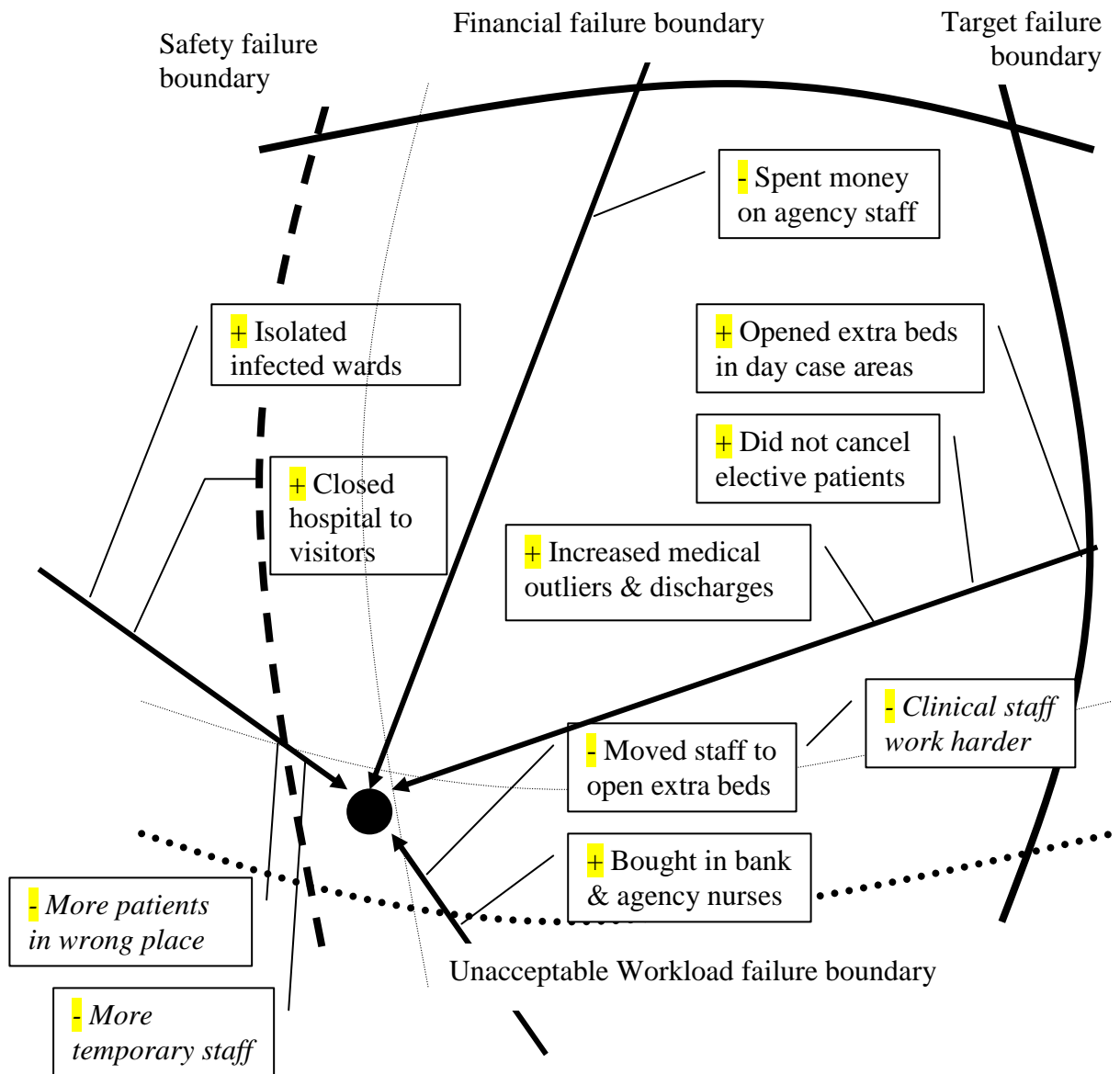


Figure 7.6: Event one – SWE (v3) with operating point and actions and consequences associated with gradients from each boundary
 (+ = compensating action to keep OP away from that boundary; - = actions or consequences that brought OP closer to boundary)

From this data there are a number of conclusions to be drawn about the dimensions being investigated. The conclusions are summarised in Table 7.2 and 7.3. In Chapter 9 these results are discussed in relation to the literature.

Construct	Dimension	Examples
Operating Point	Movement	Special cause increase in infection rate made the OP unstable. Rapid movement of OP evidenced by infection rate and the change in the number of medical outliers
Operating Point	Location	Normal number of medical outliers >20 therefore in buffer area of safety failure prior to event During virus breached safety failure boundary due to the number of patients infected and number of outliers

Table 7.2: Summary of the dimensions of the ‘operating point’ during Event 1

Construct	Dimension	Examples
Gradients	Scale	Mixed scale interactions as staff sought to respond to safety requirement to isolate wards and meet target requirements by spending money and staff working harder
Gradients	Pressure	High pressure to control infection; low pressure towards safety implications for other patients High pressure towards achieving targets Some pressure to reduce staff workload – bought in additional nurses but accepted high workload Low pressure from finance gradient

Table 7.3: Summary of the dimensions of the ‘gradients’ during Event 1

7.3 Event two – post Christmas surge in emergency admissions

There was a surge in the number of emergency admissions immediately after Christmas 2008 in CS 1. The Site Management Team used a daily predictor for the number of medical emergency patients that they expected needed admission [CS 1: 1.1; CS 2: 2.2] (see Appendix 7.1). The predictor used historical data and took account of the day of the week, time of the year and other factors. It is regarded as being very helpful and in the reported experience of the site managers, ‘pretty accurate’. Table 7.4 shows that over period of 17 days commencing the 29 December 2008 the predicted volume of emergency medical admissions underestimated the actual number. There were 186 more medical emergency admissions than expected in 17 days.

Date	Day	Admitted	Predicted	Medical Outliers
29 Dec	Mon	61	42	66
30	Tues	51	40	77
31	Wed	61	46	74
1 Jan	Thurs	39	41	No data
2	Fri	55	41	71
3	Sat	61	33	113
4	Sun	37	30	109
5	Mon	58	42	113
6	Tues	47	41	109
7	Wed	53	47	99
8	Thurs	49	41	101
9	Fri	59	41	105
10	Sat	37	33	No data
11	Sun	40	30	No data
12	Mon	51	42	95
13	Tues	55	41	100
14	Wed	50	47	107
		864	678	

Table 7.4: Medical Emergency Admissions and Medical Outliers (29th Dec 2008 – 14th Jan, 2009)

The *predicted* number of admissions was used to calculate the upper and lower control limits. The statistical process control chart (SPC) (Figure 7.7) shows that the *actual* emergency admissions breached the upper control limit (UCL) of the predicted admissions on seven out of seventeen days. On all but three days, the actual admissions are on or above the mean of the predicted admissions. This analysis indicates the presence of special cause variation from the expected level of admissions (Wheeler, 2003). From a resilience perspective, this type of variation can be described as a perturbation to the system due to the continuous stress of high levels of admissions (Wreathrall, 2006).

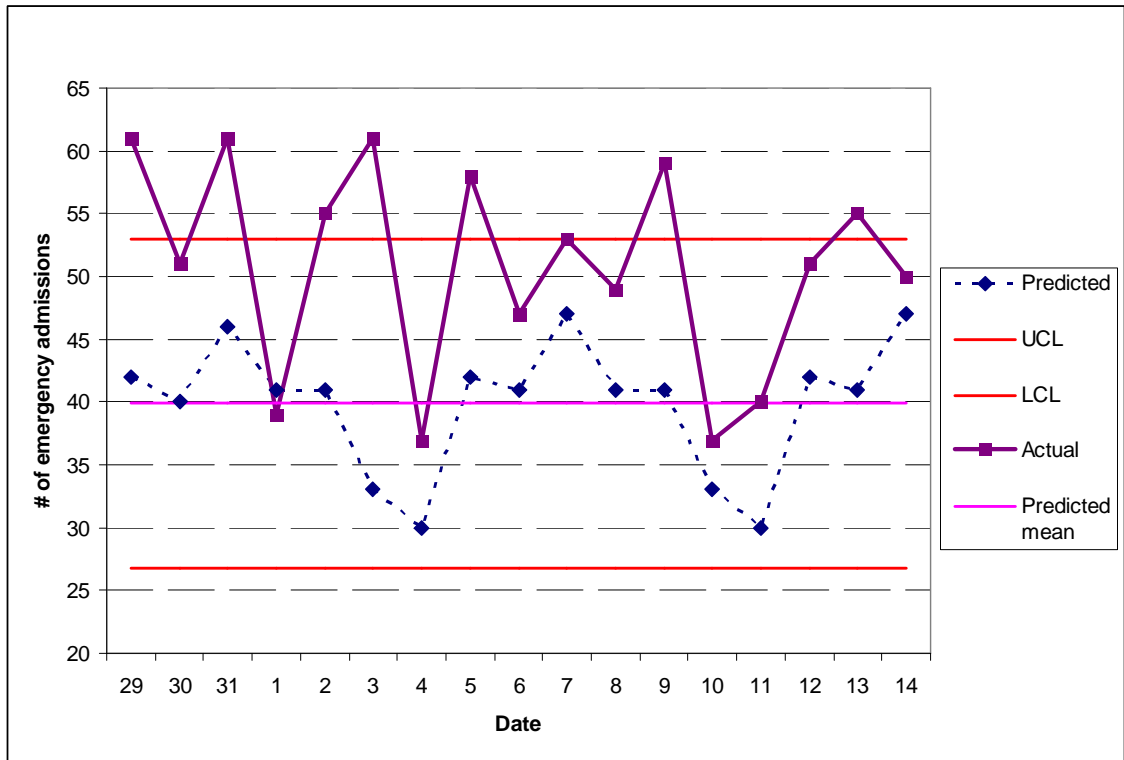


Figure 7.7: CS 1 Predicted emergency medical admission with control limits with actual admissions Dec 08 – Jan 09

The rise in demand contributed to the situation where there were over 110 medical patients on non-medical wards by Saturday 3rd January 2009 (Figure 7.8).

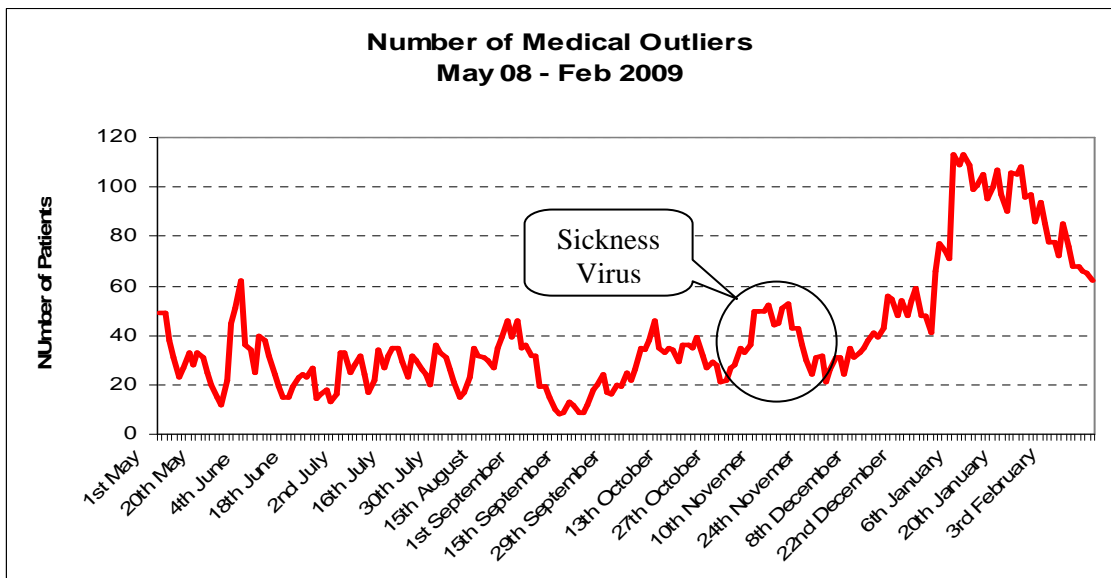


Figure 7.8: Number of medical outliers by day May 08 – Feb 09

7.3.1 Response- staff workload and patient safety

The additional capacity that had been opened during the sickness virus was again used. Daycase beds were staffed overnight for inpatients; the pre-admission areas had beds put in them and used for inpatients, and a orthopaedic and a surgical ward became medical wards. Staffing those areas was a constant problem. A medical consultant was allocated to look after the medical patients on each of the converted surgical wards, which doubled their inpatient workload [CS 1.25]. Junior medical staff were redeployed and nurses were moved from across the hospital to staff the orthopaedic ward, which had been closed for the holiday period [CS 1: SWE Staff workload].

Providing nurse staffing to look after inpatient in the pre-admission area in Orthopaedics located at the other end of the hospital from EMU was not easy. EMU had responsibility for that overflow area:

“You’ve got nurses who haven’t worked together before of course you know it’s a core of it in the EMU the other bit is made up from nurses in the medical division and bank nurses so you know they are not working together they don’t know each other necessarily and we are not having consistency with the same staff looking after the patients and the system has been put together quite quickly that ward was cobbled together two weekends ago and it wasn’t an operational area it was pre assessment area so even now we are still you know we haven’t got the right stores in place and the pharmacy’s not quite right because of course it’s not an established in-patient area.” (Senior Nurse CS 1) [CS 1: AD; Capacity]

This is an example of the unacceptable workload gradient being relaxed in terms of putting staff in unknown and unsuitable environments to work in teams that were not used to being together. The safety implications of this type of situation did not appear to be a high priority. When staff were asked about the potential issues of patient safety for medical outliers there are no direct concerns expressed. One doctor commented:

“I don’t know if you could demonstrate in any trial that they did less well than the people on a general medical ward yet it’s a pain having to go out and look after them and it creates more work but we are professional people so you just do the work in a professional way as it’s the same body of work.” (Doctor CS 1)

There was an acceptance that the continuity of care was interrupted for those patients who are moved from one ward to another. A patient moved to a ward not covered by the consultant team who were treating them, would transfer to the care of the medical team on the new ward. As a doctor commented:

“...wherever they pitch up they become somebody else’s patient. That’s one of the things that really dissatisfies everybody because there’s another team got to get to know the patient, they’ve got to go over the story again and it’s never going to be quite the same as having seen their illness evolve or hopefully resolve.” (Doctor CS 1)

On one day it was observed on two separate occasions that a junior doctor finished writing up the medical notes following his assessment of the patient on EMU, turned around and found that the patient had been moved to another ward. He shrugged his shoulders and moved onto his next task [FN 2.39]. The patients had been transferred to another ward without their medical notes, which is the key means of communication about their investigations and treatment. Conceptually, this can be seen as an example of the gradient ‘towards increased production’ exerting pressure to keep the patients flowing through the hospital. The willingness to move patients without their medical notes suggests that the need to achieve the productivity requirements had a higher priority than the potential safety of the patient. Conceptually, the marginal zone boundary was shifted outwards as it became acceptable practice to move patients without their medical notes. Further evidence of the boundary shifting is explored further in Section 7.3.3.

7.3.2 Response – decision making hierarchy

Given the scale of the capacity problems an internal ‘major incident’ was initiated. A meeting led by the Chief Executive developed a decision making hierarchy for prioritising admissions. Emergency admissions had to be admitted and had to be ‘admitted well and dealt with as well as we possibly can given the demand that we’ve got’ (CEO). The second priority was clinically urgent patients and

“...then actually nothing else matters because that’s what we’ve got to achieve so if that means we cancel elective surgery we cancel orthopaedics, whilst it’s unpleasant for the patient that’s been cancelled they are actually not going to come to significant harm. If we cancel somebody’s varicose veins it’s unpleasant but they are not going to come to significant harm.” (CEO CS 1)

The hospital was ahead of the RTT target in most specialities except orthopaedics. This provided some flexibility in identifying non-clinically urgent patients to be cancelled or treated elsewhere without creating major target issues. Unlike during the sickness virus, senior managers took the decision to relax the pressure on the system by cancelling some elective admissions.

At certain points during this period it was noticeable how involved senior managers and Directors became in the micro management of patient movements. On several occasions Senior Nurses, Managers and Directors acted as porters in the late afternoon, early evening and even late at night to help move patients from the ED to EMU or from EMU to other wards. A ward sister was observed receiving a phone call from a manager who asked her to admit a patient from ED into the treatment room until a bed became available [FN 2.58]. The sister refused and expressed the view that to admit a patient into a treatment room was a line she was not prepared to cross. The site manager, who also observed the phone call, was annoyed as she had a plan in place for that patient. This is an example of what might be described as a collapsing hierarchy where senior managers intervene in the work of more junior staff in an effort to resolve a problem that is often already being addressed. Conceptually, it can be seen as a middle manager placing more importance on keeping the OP away from the target failure boundary than the patient safety or staff workload boundaries.

7.3.3 Consequence – shifting the safety failure marginal zone boundary

On the afternoon of Friday 2nd January 2009 it became apparent that the number of emergency admissions was going to be larger than expected and there were not enough empty beds [FN 2.26]. The rich data, captured during a short period during event two, shows how the ‘drift to danger’ can be accelerated as the marginal zone patient safety boundary is shifted due to the pressure to maintain production (Woods et al., 2007, Snook, 2000).

The Lead Nurse for the Control of Infection had given advice at a previous bed meeting on the 29th December. She had stated clearly that medical patients should not be outlied on the gynaecology ward [FN 2.23]. The reason is that medical patients are at higher risk of having the Norovirus. The doctors who work on the gynaecology ward also work in the maternity unit helping to deliver new born babies. Therefore, if the Norovirus was to become present on the gynaecology ward, there was a potential risk of it spreading to the new born children and their mothers with severe consequences. Conceptually, the marginal zone boundary location had been clearly articulated by the Lead Nurse for Infection Control.

In addition to the infection control concerns the skills and experience of staff on the gynaecology ward was also an issue. The gynaecology ward deals mainly with female patients requiring surgery, who are otherwise well, and only stay for a few days. The staff do not expect to deal with complex medical patients and are not familiar with many of the drugs used for such patients. Up to this point in early January, when patients had to be outlied, female surgical patients had been selected to be transferred to the gynaecology ward [FN 2.23; 32].

On Friday the 2nd January 2009, there had already been an extra 3pm bed meeting. It was reported that the number of medical outliers was up to 71 and the anticipated emergency admissions was 10 more than the expected. The suggestion was made that the empty beds on the gynaecology ward would have to be used [FN 2.34]. It became clear that the surgical unit had already moved all those patients who were suitable to move. The view was voiced that “there may be medical patients who could move to the gynaecology ward”. It was decided to review the situation at the 4pm meeting.

At the same time there was concern about staffing numbers for the weekend. Requests had been sent to agencies but there had been no response. The Lead Nurse for Medicine then asked permission to go outside the approved NHS contract for agency staff and contact a more expensive company that was not on the approved list. The Lead Nurse stated: “We won’t get quality but we can put them in areas and move others.” Permission was granted by the Director of Operations [FN 2.34]. This is an example of the financial (spend extra money) and patient safety (accept poorer quality staff) gradients being relaxed in an attempt to ease the workload for nursing staff. The action was also motivated by the needs to keep wards open generated by the gradient towards production.

At the 4pm meeting the staff attending overflowed into the corridor. The CEO and Medical Director were present and the Director of Operations (DoO) ran the meeting. The hospital expected 18 more patients than they had beds for. The DoO asked the Control of Infection nurse who was present (not the Lead Nurse) for his view about moving medical patients onto the gynaecology ward. He said that “if there was no alternative then it would be OK”. The Medical Director, who was standing just outside the door in the corridor, then suggested some criteria about selecting medical patients for the gynaecology ward: “avoid elderly and those on anti-biotics” The response of

those inside the room was laughter and someone said: “There are no young patients” [FN 2.35].

The Lead Nurse for Medicine was observed at this time as she spent a number of hours micro managing the flow of patients from ED and EMU [FN 2.38]. After the 4pm bed meeting the Lead Nurse went directly to EMU. She asked the Matron to identify 6 patients for transfer to the gynaecology ward; no mention was made of any selection criteria. Once the patients were identified, the Lead Nurse and a Senior Matron moved the patients, their belongings and notes to the gynaecology ward. One of the patients transferred was a cardiology patient. The Lead Nurse had to collect the required drugs for that patient from Pharmacy and take them to the gynaecology ward. She assumed the nurse knew how to administer the drugs. She told the nurse that other medical drugs would be available from a medical ward if required [FN 2.38].

Once the ‘no medical patients on the gynaecology ward’ line had been crossed, the following day (Saturday), when again the emergency admissions were much higher than expected, over twenty medical patients were transferred to that ward. When the Lead Nurse for Infection Control returned from leave the following Tuesday, she inquired who had made the decision stating that “You will live to regret that decision” [FN 2.66]. The Chief Operating Officer’s view of the decision was:

“I think it was the right decision to make at the time but what we didn’t do was safeguard how those beds were going to be managed over the weekend and you know we should have and in fact I’ve been sought by the control of infection lead quite rightly and taken to task over it.” (COO CS1)

There was a subsequent series of conversations and the criteria for selecting patients to move to the gynaecology ward was tightened to only surgical inpatients or medical female day case patients. Follow up information shows that the gynaecology ward has now been re-designated as a ‘woman’s surgical ward’.

7.3.4 Analysis of event two

The rich data gathered over of a few hours within the hospital describes the multiple factors that were involved as medical patients overflowed into all parts of the hospital. As demand outstripped the supply of empty beds the situation produced a response from decision makers in the hospital. In this event, as the cumulative volume of emergency admissions occupied the inpatient beds, there was a willingness to cancel selected clinically non-urgent elective admissions. It also illustrates the moment when the reality

of the situation, (more patients than empty beds in the right areas), forced managers to take a decision that in conceptual terms, shifted the position of the marginal zone boundary for patient safety towards accepting higher risks (Figure 7.9). Managers felt that they had no other option other than to ‘trade off’ (Hollnagel, 2009a) the risks of using the gynaecology ward with the risk of not being able to admit patients from the ED.

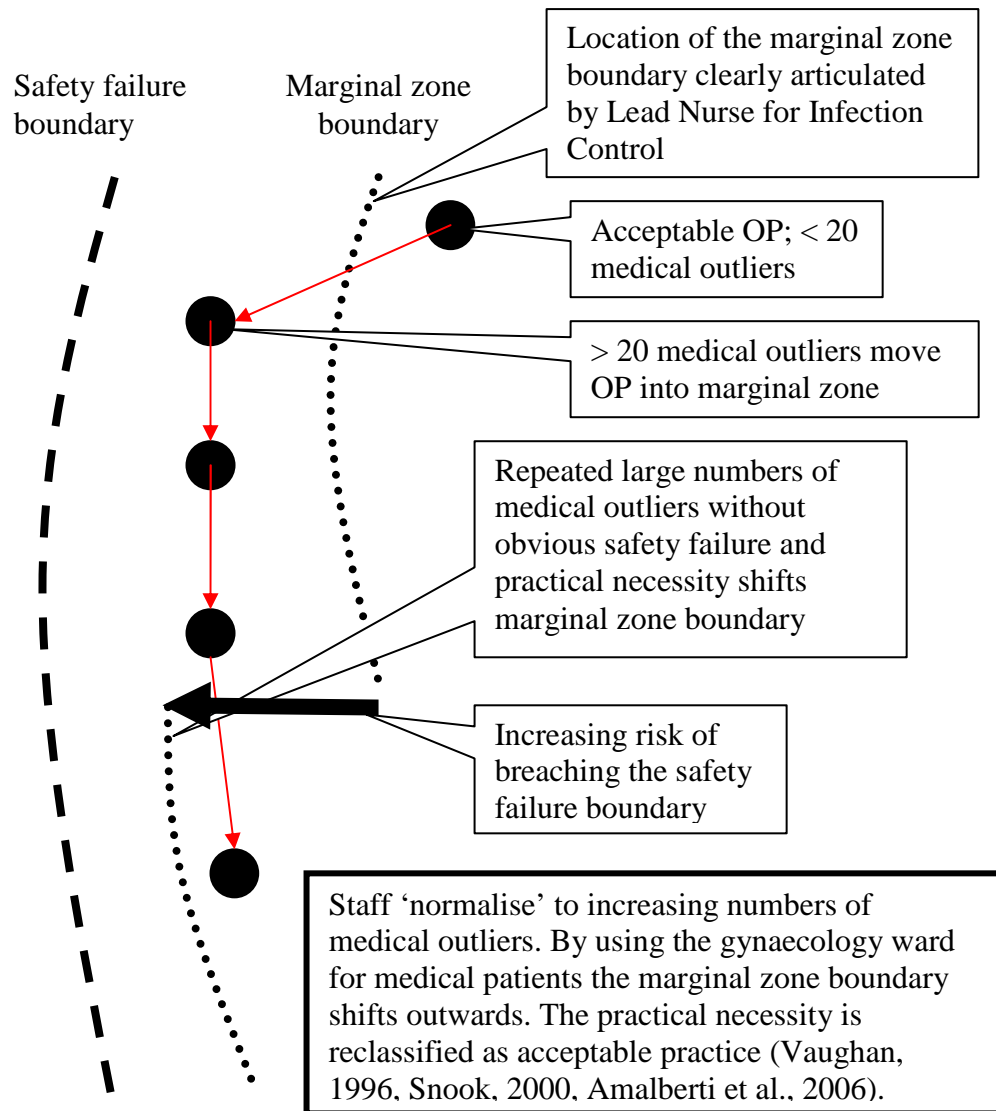


Figure 7.9: CS 1 Shifting the marginal zone boundary outwards

The pattern of response to the situation is similar to the sickness virus. During this event there was some relaxation of the elective RTT in some surgical specialities so that the emergency patients could be admitted in a timely manner. The financial gradient was also relaxed to pay for additional staff. However, the need to open extra beds meant that

the staff workload increased and the number of medical outliers increased substantially. At points during the day when there were peaks in admissions the hospital system went 'solid' (Cook and Rasmussen, 2005). There were no empty beds and no admissions were possible until a patient was discharged. Senior managers became involved in the micro management of patient movements [FN 2.56, 58].

Whilst staff claimed that patient safety was their top priority, the evidence from the observations is that patient flow was their highest priority. Such evidence includes using the gynaecology ward to keep the flow of medical patients moving to allow admissions from ED. The assumption appears to be that it was 'safer' to have high numbers of medical outliers than not to be able to admit emergency and clinically urgent elective patients. However, the data does suggest that such a view is influenced by the need to achieve the waiting time targets.

The situation over this period presented decision makers with difficult choices. There were no simple answers and under considerable pressure managers and clinicians sought to make the least worst choices. As no immediate harm appears to have occurred to medical patients who were outlied to the gynaecology ward, staff normalised very quickly to that practice. The Lead Nurse for the Control of Infection challenged that practice. It can be inferred that her job was focused on not shifting or breaching the patient safety boundary in relation to infections.

The data display for this situation is shown in Figure 7.10:

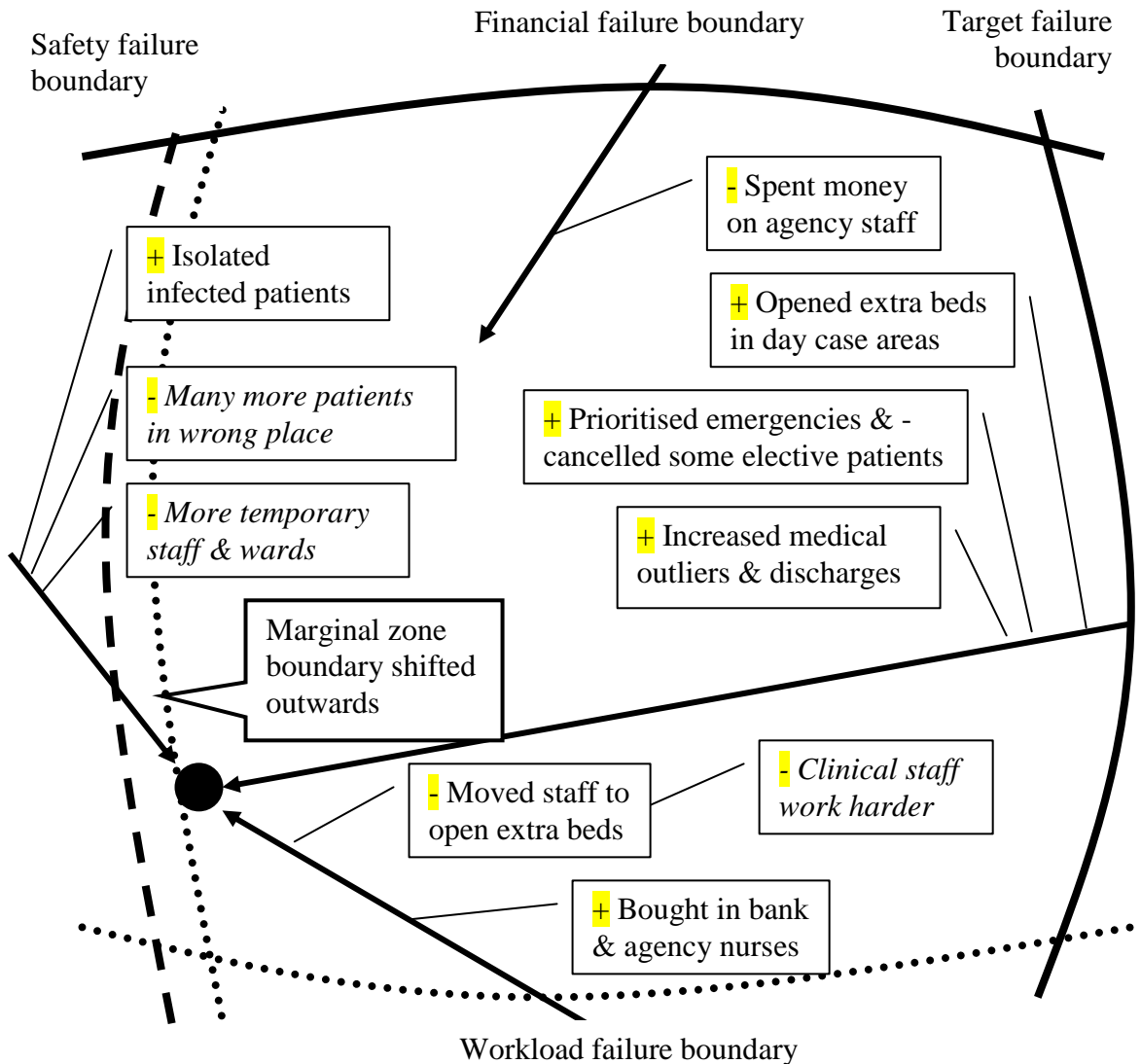


Figure 7.10: Event two – SWE (v3) with operating point and actions and consequences associated with gradients from each boundary
 (+ = compensating action to keep OP away from that boundary; - = actions or consequences that brought OP closer to boundary)

The dimensions during the surge in demand are summarised in Tables 7.5 and 7.6:

Construct	Dimension	Examples
Operating Point (Event 2)	Movement	Special cause increase in medical emergencies made the OP unstable Rapid change in the number of medical outliers created rapid movement of the OP
Operating Point (Event 2)	Location	Normal number of medical outliers >20 therefore in marginal zone of safety failure prior to event During event breached safety failure boundary due to wards used and number of outliers

Table 7.5: Summary of the dimensions of the ‘operating point’ during the Event 2

Construct	Dimension	Examples
Gradients (Event 2)	Scale	Mixed scale interactions as staff sought to respond to safety requirement to admit emergency patients and meet ED target requirements. Spent extra money on staffing, staff working harder and cancelled some RTT patients
Gradients (Event 2)	Pressure	High pressure to admit emergencies; lower pressure to safety implications for patients outlied High pressure due to ED target; flexible use of RTT target Some pressure to reduce staff workload – bought in additional nurses but accepted high workload and temporary ward environments Low pressure from finance – spent money on staffing.

Table 7.6: Summary of the dimensions of the ‘gradients’ during Event 2

7.4 Event three – the flow of emergency patients on one day

The previous two events took place over a number of days. Data from CS 2 confirms a similar pattern of response to infection and peaks in demand. However, the smaller size of CS 2 meant that the medical outlier situation was more easily managed by both clinical and management staff. Therefore, the OP returned to within the envelope more quickly than in CS 1. Event three from CS 2 is a description of how emergency patients were managed during one day [FN 4.1-20]. This is used to illustrate how the capacity of the hospital to manage the admissions over a short period of time can impact on the patients. Observations were made of the Medical Assessment Unit (MAU) (similar to the EMU in CS 1) and the Emergency Department (ED) during the 29th December 2009. Interviewees were conducted to clarify and triangulate the observational data. Access to this hospital was limited and this day was chosen to replicate a similar observation exercise on the same day the previous year in CS 1.

7.4.1 The flow of patients

The consultant ward round on MAU started at 8.30am. Fourteen medical emergency patients had been admitted overnight. There was one empty bed on MAU. One, and later two bays, had mixed sex patients (breach of single sex target). This breach was regarded as a consequence of prioritising patient care:

“Now I think the executive made a clear decision last year earlier this year that, no last year sorry, it’s January, that patients’ care was not going to suffer as a result of the single sex requirement and largely that’s done but there are still times when you know an inordinate amount of shuffling beds just to fix one kind of problem.” (Doctor CS 2) [CS 2: SWE; Targets]

The first patients seen by the consultant on the ward round were those admitted by the junior doctors who had been on the night shift. They ‘handed over’ their patients before going off duty. The patients admitted the previous evening by the late shift of junior doctors were reviewed later; on that day starting at 11.10am.

The consultant reviewed each patient and decided on their treatment plan and likely length of stay. The nurse running the MAU accompanied the doctors and co-ordinated future actions with the junior doctors on the day shift. The nurse reminded the doctor that ‘there are no medical beds’ in the hospital so for him to consider alternative options to admission where that was feasible [FN 4.8]. By 9.50am there were two patients in the ED who needed to be admitted to MAU. The bed manager had found three beds in surgical wards for MAU patients. The nurse asked the consultant to identify any patients for admission who were suitable to go straight to a surgical ward. A patient who was waiting for an MRI scan, but otherwise had recovered from her acute episode of illness, was deemed suitable to move to an orthopaedic bed.

At the 12.30pm bed meeting it was reported that the number of emergency admissions on the previous Sunday had been exceptionally high and that due to many doctors being on holiday, the number of discharges was lower than expected [FN 4.12]. There was no visual display of bed occupancy. The predictor of admissions for that day was e-mailed to managers [CS 2: xls 2.2]. The on-call manager was present at the meeting and asked why the patient waiting for the MRI could not have been kept in a chair on MAU rather than occupying a surgical bed. The on-call manager took the view that “Now we have 21 outliers we have lost the plot.” Medical teams were to be ‘chased’ to see the outliers [FN 4.14].

By 2pm the ED was full. There were medical patients referred by their GP in the ED waiting room because the MAU was full. Two patients were in the resuscitation room inappropriately (they were not seriously ill but there were no other trolley spaces available). If a seriously ill patient arrived the resuscitation room had to be available for life saving treatment. In this situation it was full. As one doctor commented:

“...there are times when every bed in the department is occupied and that clearly is you know potentially catastrophic if somebody comes in in cardiac arrest or a major trauma and there isn't a bed to put them on and you know the reality is that in that situation somebody would be moved and the patient would be taken through to resus on an ambulance gurney and treated on that but you know that is clearly a very dangerous situation.” (ED Doctor CS 2) [CS 2: SWE; Patient safety]

There were five patients on ambulance trolleys waiting in the corridor [FN 4.18]. There was a social worker available to ED who could fast track services for patients in the community to avoid hospital admission. She was unavailable for the next hour as she caught up with work done to help MAU. A large white board on the wall was kept updated by an administrator. This provided patient details including the time they would breach the four hour target.

The Sister in charge of ED spent the next few hours micro managing the flow of patients [FN 4.18, 20]. When a patient in the minors section of the ED was discharged she moved a child and parents from the paediatric room into the vacant adult cubicle. Then she moved the medical patient from the waiting room into the paediatric room and got a doctor to assess the patient and then transfer them to MAU. In the meantime a paediatrician arrived and could not find her patient. The Sister assessed the ambulance trolley patients and sent one off for an X-ray. An extra trolley was found and put in the resuscitation room and one of the ambulance trolley patients was moved in there. The social worker arrived earlier than expected and arranged the discharge of a patient. The MAU were able to transfer some of their patients into surgical beds. The ED was then able to transfer a number of patients to MAU and the pressure within ED reduced. The Sister insisted that the first priority was patient safety and targets came second. The constant juggling of resources and moving patients did mean that staff were working hard. The four hour target influenced how resources were deployed in the department. The dynamics that occurred in the ED was very similar to the whole hospital system. In the ED those dynamics were more visible and occurred in a tight time frame. As one doctor commented:

“I think what happens is that there is a slightly blurred tipping point where the department makes a transition from being busy to becoming I guess not quite out of control but certainly very inefficient and if I can give you examples of that to sort of bring to light what I mean. Often as a doctor you will see a patient and you will want things done for the patient to assist in making a diagnosis to start treatment or whatever and you need members of the nursing staff to execute those actions. When the department gets very busy it can be very difficult to find the correct member of nursing staff or even a member of nursing staff who is able to attend to those jobs and so you can waste time in trying to find somebody and obviously because they are busy it takes longer for a job to be done and so a patient who might potentially be moved on through the department either to go home or to be admitted quite quickly will stay much longer and because they are there and their bed space is occupied the department can't flow and there is as I say a tipping point when the department is sort of probably 90% full when all of a sudden it will go to 100% full and there will be a queue of ambulance trolleys in the corridor simply because we've reached a sort of gridlock situation... Basically what happens is that people work harder and I think they work harder and the nurses don't get their meal breaks but they don't work more effectively and there is no pressure release valve down here...” (ED Doctor CS 2) [CS 2: AD; Flow]

In conceptual terms it can be argued that the ‘tipping point’ referred to in this quote is when the OP moves into the patient safety failure boundary marginal zone and the ED system goes ‘solid’ (no empty trolleys). In such a situation the ED as a system becomes tightly coupled and therefore, vulnerable to the OP breaching the boundaries of the SWE in any one of four directions. When the system becomes ‘very inefficient’ it is argued that the dynamics of the situation have changed due to the increased number of reinforcing feedback loops between the parts that previously had not been coupled together. The task of managing the situation becomes more complex. In Chapter 8 the changes in the dynamics are illustrated using CLDs.

7.4.2 Implications for patients

From the description of the flow of emergency patients there are a number of observations about the system behaviour and the implications for patients. It would appear that both departments were heavily dependent on the wider hospital to have empty beds. When there was a delay in accessing empty beds, the flow of patients through the hospital was delayed, and the OP moved closer if not through the safety failure and unacceptable workload boundaries.

This data from ED and MAU suggests that the OP can move very quickly to a point where staff are overworked and patient safety is potentially compromised. However, both the ED and MAU are to a large extent dependent on the rest of the hospital system

to provide them with empty beds in the right areas in a timely manner. When beds are not available then the system becomes more tightly coupled with potential consequences.

The data concerning a particular patient provides an insight about how the dynamics and interactions within the hospital system can impact on the outcome for individuals. In this incident, told by an ED doctor in CS 2 who investigated the situation, the patient does breach the four hour target in ED. However, it is a combination of the workload of the MAU, the inexperience of staff, poor handover of information and the difficulty in accessing a single room that created a poor outcome.

“He was an elderly man who was admitted with atria fibrillation and diarrhoea and he came to the emergency department having been referred by a general practitioner but he came to the emergency department because the medical assessment unit was full and he came into one of our bays and the medical team on duty were told about him but were busy on the medical assessment unit and didn't get to see him. He arrived at around four o'clock in the afternoon and his wife who was with him was very unhappy for him really quite early on with the delay in seeing him although he was reasonably stable he certainly would have benefited from earlier treatment and he finally left to go to the ward at about ten o'clock in the evening. He would have been free to go to the medical assessment unit at about eight as I understand it but because he had diarrhoea he needed a single room and the decision was made that he should wait and go to a single room on the ward rather than leave the single room down here and so it was about ten o'clock when he got to the ward which was just around the time that doctors handover and so he didn't see a doctor until after the handover period and so the person who had spoken to the GP wasn't the person who saw him and his treatment started then and certainly his treatment from that junior doctor was not as good as it could have been and the choice of drugs for atria fibrillation I would say was clearly wrong and it wasn't corrected until the following morning because that was the first time he was seen by a senior. If he had been handled seamlessly he would have gone to the medical assessment unit at four o'clock in the afternoon, would have been picked up on the consultant ward round at five thirty, started on the appropriate drugs at six o'clock and he would have been better as a result. The eventual outcome for him was that he died.” (Doctor CS 2) [CS 2: SWE; Patient safety]

There were a number of influences on the OP by the gradients working at different levels within this incident. At the micro level the patient breached the four hour target. Due to his diarrhoea, he was deemed to be an infection risk and was therefore placed in a single room in the ED rather than being transferred directly to the MAU for treatment. It would appear that the patient safety concern about infection was prioritised over his need for his heart condition to be treated. In maintaining the patient safety requirement to isolate the patient, a number of less explicit safety issues arose; delayed treatment of

serious underlying heart condition, poor handover of information, questionable prescribing and a delay in being reviewed by a senior doctor.

At the meso departmental level the lack of capacity within MAU to accept an 'infectious' patient meant that he was held in ED. The patient was not being treated by the ED doctors, because he was designated a GP admission. He was the responsibility of the specialist medical doctors on MAU who were busy looking after a ward already full of patients. The patient was therefore in a queue for MAU receiving nursing care but not medical treatment in the ED. The design of the admission system, the supervision of junior staff, the handover of information and the timing of senior review do not appear as important patient safety priorities in comparison to isolating potentially infectious patients.

At the macro hospital level the reason that MAU could not accept the patient was because they could not move a patient from a side room to another suitable bed elsewhere in the hospital. Following high profile patient safety failures where organisations were found to be lax in controlling infections (Healthcare Commission, 2006b), the Department of Health imposed targets to reduce health care acquired infections (Department of Health, 2007). Hospitals therefore prioritised isolating potentially infectious patients. Hospital buildings are not designed to have large numbers of single rooms and when a hospital is operating at over 90% capacity there is little flexibility to cope with peaks in the number of patients with infections.

The European Working Time Directive influences the way hospitals deploy doctors. Most junior doctors work a shift system to limit their hours of work. The implications are that the junior staff are not working with the same senior doctors on each shift and the number of handovers of patients at the change of shift has increased. In the description of the ward round on MAU, the consultant started with the patients admitted by the overnight shift of doctors, to allow them to go home. The priority order of reviewing the patients was not based on clinical need but keeping the junior doctors working hours within the required limit. From a conceptual perspective this can be viewed as an 'upward scale' action to keep the OP away from the unacceptable workload boundary for junior doctors. The interrelationship of that action on the other gradients, do not appear to have been anticipated. For example, the delay in the consultant reviewing a seriously ill patient due to prioritising the time of admission over

clinical need.

A key point to draw out from the analysis of this incident is that the gradients create influences often for the best of reasons. These include reducing healthcare acquired infections and the working hours of doctors. However, such gradients when combined with other influences, such as waiting time and infection control targets, can create conditions that impact the treatment of individual patients. Often the impacts can be positive; patients are kept safe from infection, admitted more quickly and are treated by doctors who are not over tired. Yet, it would appear that a system can equally be vulnerable to failures associated with not paying sufficient attention to the wider dynamics created by prioritising the isolation of infected patients, overcrowding the hospital to meet waiting time targets and reducing the hours of work for doctors. The dimensions of the OP and gradients constructs are summarised with examples from Event 3 in Tables 7.7 and 7.8.

Construct	Dimension	Examples
Operating Point (Event 3)	Movement	With high levels of occupancy the ED ‘tips’ into situation of inefficiency. The OP became unstable and could breach any boundary.
Operating Point (Event 3)	Location	The lack of capacity to isolate infectious patients in a timely manner located the OP within the marginal patient safety zone.

Table 7.7: Summary of the dimensions of the ‘operating point’ during Event 3

Construct	Dimension	Examples
Gradients (Event 3)	Scale	The European Working Time Agreement exerted downward scale gradient on the hours doctors work. Upward scale gradient in how ward rounds were conducted placing time of admission before clinical need to reduce junior doctors hours of work. National targets on infection control exerted downward scale pressure to isolate infected patients. Upward scale gradient of prioritising control of infection over other clinical needs of patient held in ED.
Gradients (Event 3)	Pressure	The pressure to meet multiple waiting time targets increases the hospital occupancy making it harder to admit patients from ED in a timely manner.

Table 7.8: Summary of the dimensions of the gradients during Event 3

7.5 Summary

It is recognised that the constructs of the OP and gradients are closely linked to the dimensions of the boundaries discussed in Chapter 6. There are however, certain dimensions of the OP and gradients which can be inferred from the case study data presented above. Many of the conceptual points relating to the SWE (v3) model have been made in the analysis of each of the events. A brief summary of the key points are set out in the next two sections.

7.5.1 Operating Point

The OP depicts the working conditions of the system in relation to the performance failure boundaries. A key idea from the SWE (v3) model is that a resilient system is one that can remain operating within the envelope in the face of perturbations or continuous stress. The two dimensions of the OP construct, 'movement' and 'location' are derived from the literature and populated from the case study data. Cook and Rasmussen (2005) argue that it is these two dimensions that indicate the 'reliability' of the system. In the context of the model the 'reliability' is about the ability of the system to keep the OP within the envelope.

The 'movement' dimension of the OP can be seen as representing the operating conditions of the system. When a system faces a disturbance there is the potential for the OP to move from a stable to unstable state (Taylor and Ford, 2006, Holling, 1973). Rapid and large movements are symptoms of low reliability (Cook and Rasmussen, 2005). Rapid movement of the OP indicates problems of stability and reduces the time available to take compensating actions to avoid breaching a boundary. However, slow movement of the OP is not always safe. It can indicate a 'drift to danger' where new norms of performance are accepted over time and a breach of a boundary creeps up on decision makers (Snook, 2000, Vaughan, 1996, Amalberti et al., 2006).

Closely linked to the 'movement' is the 'location' dimension of the OP in relation to the boundaries. The literature suggests that the cost of keeping the OP well away from the safety failure boundary is unsustainable (Reason, 1997, Cook and Rasmussen, 2005, Flin et al., 2008). The consequence is that, in most systems, the OP is usually located close to or in the marginal zone of the safety failure boundary. In reality there is no visible OP. Therefore, observations of what processes are in place to monitor

performance in relation to the failure boundaries are used to provide an insight into the location and movement of the OP.

In this research the number of infected patients (Event 1) and the number of medical outliers (Event 1 and 2) are used as proxy measures for the location of the OP. When there are high numbers of medical outliers, which indicates overcrowding of the bed capacity, the location is conceived to be close to or breaching the patient safety failure boundary. These assumptions are based on the literature which indicates an association of higher mortality and patient harm with overcrowded hospitals (Cameron, 2006, Fatovich et al., 2005, Richardson, 2006, Sprivulis et al., 2006, Trzeciak and Rivers, 2003). The data from the case studies suggests that staff regard medical outliers as far from ideal, but it is inferred that they became an accepted violation (Amalberti et al., 2006). The level of harm generated by having medical outliers is not assessed nor visible to staff. Therefore, the benefit of having medical outliers is that the hospital can admit the very visible emergency and elective patients. This benefit is thought by staff to outweigh the largely unseen cost to patients of being treated in the wrong place.

7.5.2 Gradients

The gradients in the original SWE model are depicted as linear influences on the OP (Rasmussen, 1997). However, as noted previously, there is a dynamic set of interrelationships that occurs. Conceptually, the gradients can be regarded as influences derived from the 'social structure'. However, within a system the 'agents' respond to the influences of the gradients. The responses in turn can change the influence exerted by the gradient. The two dimensions of 'scale' and 'pressure' identified from the literature are investigated. These two dimensions each inform the other and are therefore regarded as different lenses on the same construct.

Woods (2006) suggests that there are 'scale interactions' that create dynamic influences on the OP. He argues that there is a 'downward scale' influence that comes from the 'blunt' end of the system. At the same time there is an 'upward scale' influence generated by those at the 'sharp end' of the system as they work in the context of competing downward influences. The influences on the OP from the gradients are therefore a complex set of interacting themes.

A means of assessing the patterns of influence is to examine the 'pressure' from the

gradients on the OP through the actions and opinions of staff, and the policy requirements set out by the wider 'social structure'. There does appear to be a conceptual link between the 'visibility' of a boundary, examined in Chapter 6, and the 'pressure' creating the need to respond to the gradient relating to that boundary. The clear examples are control of infection and the four hour ED target. There was strong 'downward scale' pressure from the gradients for both of these issues resulting from their associated boundaries having high 'visibility'. A pattern that appears was that the control of infection requirements was the dominant gradient when dealing with infected patients in real time. However, from CS 1 it is noted that when decision makers face the situation of deciding between future infection risks and the need to create capacity to manage the rate of admissions by outlying patient on the gynaecology ward, the gradient related to targets appears to exert the greatest pressure in the short term. Other aspects of the patient safety gradient do not appear to have the same degree of visibility or immediacy. Therefore, they do not generate the same level of 'pressure' as infection control issues. Rather than the SWE (v3) model having just one gradient from each boundary, it is possible to argue that the influences on the OP are much more nuanced and numerous.

The perception of 'pressure' in the eyes of the decision making agents influences the subsequent actions taken. Such actions can be planned, as is seen in the response to an outbreak of the sickness virus. Actions can also be reactive and generate their own momentum, as seen with the overflow of patients into the gynaecology ward.

The data presented in this chapter shows that there are a number of interacting influences on the OP which provide insights into the characteristics of CS 1 and 2. The gradient keeping the OP away from the patient safety failure boundary is largely limited to the control of infection. Other safety issues were less visible. The need to accommodate both emergency and elective admissions created large numbers of medical outliers as the medical beds became full. It is argued that such a situation conceptually pushed the OP close to if not through the patient safety boundary. As no obvious safety breach was observed, staff appeared to normalise to the practical necessity of the situation of having large numbers of outliers. Without the action to create medical outliers, then breaches in the target failure boundary would occur. Such breaches would be very visible in comparison to most patient safety breaches. The external and internal performance management and reward system reinforced the

acceptance of changes in practice to meet the practical requirements generated by the gradient towards production.

The next chapter will examine the 'structure' and 'feedback' loop constructs of the SWE (v3) model. In Chapter 9 there is a fuller discussion exploring the characteristics of the hospitals derived from the constructs of the SWE (v3) model.

Chapter 8 – Investigating the Structure and Feedback

8.1 Introduction

One of the weaknesses of the Rasmussen (1997) SWE model is that it does not fully take account of the internal dynamics that are created by the flow of work through the system. The dynamics between demand, capacity and decision makers in working with competing pressures (gradients) and constraints (boundaries) need to be included (Figure 8.1). It is argued in Chapter 5 that using SD to compliment the SWE, those dynamics can be included in the analysis.

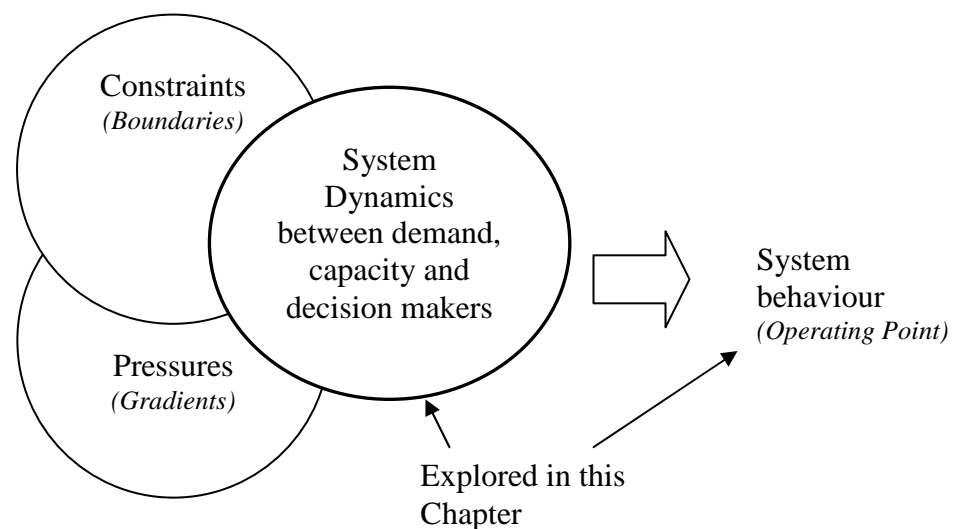


Figure 8.1: Combination and interaction of construct sets depicted by the SWE model

This chapter explores the ‘structure’ and ‘feedback’ constructs from the SWE (v3) model detailed in Chapter 5. These two constructs are used to gain insights into the dynamics that occur inside the SWE (v3), which influence the location and movement of the OP. SD provides a way to investigate the ‘structure’, which includes the ‘feedback’ loops that occur in the case study hospital systems. As noted, ‘structure’ ‘consists of the feedback loops, stocks and flows, and nonlinearities created by the interaction of the physical and institutional structure of the system with the decision-making processes of the agents acting within it’ (Sterman, 2000). In a hospital a ‘stock’ depicts the place where patients accumulate, such as a ward or department. The ‘flow’

depicts the direction of movement between stocks. 'Feedback' depicts the interrelationships between the parts of a system.

A Stock Flow Diagram (SFD) is used in Section 8.2 to illustrate the planned design of the patient flow in and out of CS 1 and 2 hospitals. The planned design is then amended in Section 8.3 to reflect the reality of the situation when the hospitals face the continuous stress of high levels of demand for inpatient beds. CLDs are used to show the feedback loops that result from the increased coupling between the parts of the system when the stocks are full and the direction of flows change. From the analysis of the data presented in the diagrams the conceptual dimensions of the 'structure' and 'feedback' constructs are identified.

The first dimension of the 'structure' construct is the 'coupling' of the parts. This depicts how closely linked together the different parts of a system are. The second dimension is the 'buffer capacity' of the stocks to be able to accommodate the variation in rates of flow into and out of the system. The dimension of the 'feedback' loops is the 'type' as to whether the loop is 'reinforcing' or 'balancing'. The implications for the 'movement' and 'location' of the OP of the model are discussed.

8.2 The planned design of the structure

The planned design of the patient flow through a hospital is partly constrained by the physical layout of the buildings. The way in which staff are deployed to meet the different requirements of patients is also a contributing factor to the design. NHS Hospitals tend to organise themselves to manage two types of patients; emergency and elective, across a range of specialities. The second major differentiation of patients is the split between medical and surgical specialities. This research concentrates at the higher level of design by considering the stocks and flows of the emergency / elective and medical / surgical work. The paediatric, obstetrics and cancer services have been excluded from the SFDs shown below. Medical specialities have a high level of emergency admissions and low numbers of elective patients [CS 1: xls 1.2]. Surgical specialities generally have more elective than emergency admissions [CS 1: xls 1.2]. Hospitals monitor the actual number of emergency and elective admissions against the planned number that has been agreed with the local Primary Care Trust (PCT) in the contract. For example, the Performance Report from CS 1 provided the Board with

graphs showing the actual number of elective and emergency admissions compared to the planned number (Figures 8.2 and 8.3).

The graphs indicate that during the period of the study in CS 1 (Oct 08 – Feb 09) there were fewer elective cases and more emergency cases than expected. In Chapter 7 it is noted that some elective cases were cancelled to provide bed capacity for the unexpected peak in emergency admission in late December 08 and early January 09. The impact of this peak in emergency demand on the ‘structure’ is shown in the next sections.

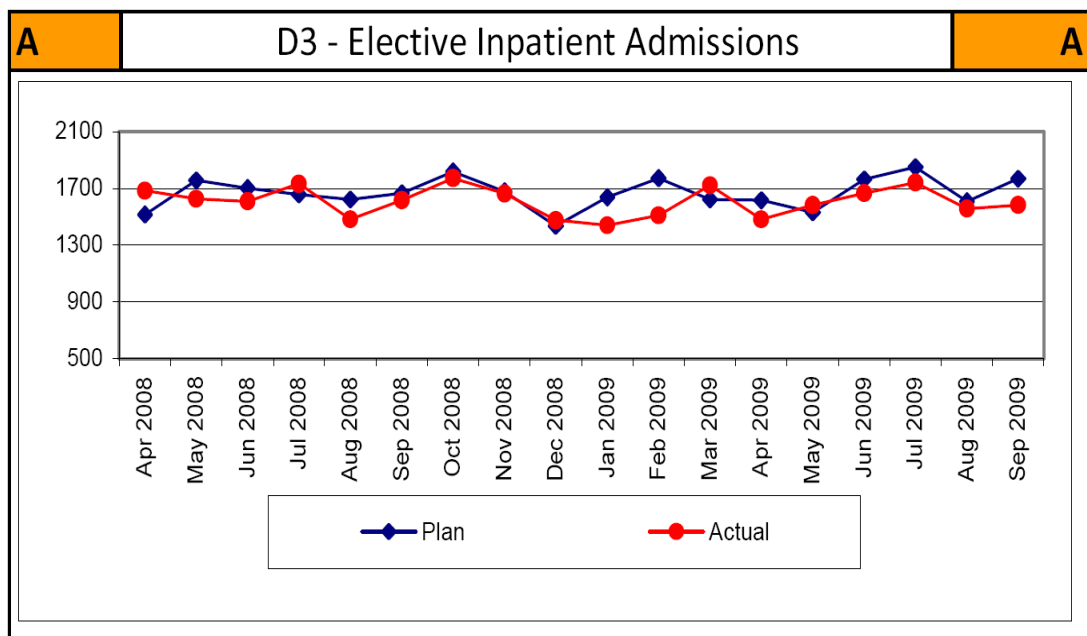


Figure 8.2: CS 1 Elective Inpatient Admissions, Plan vs Actual, Apr 08 – Sept 09 (extract from CS 1 Trust Board Performance Report, September 2009)

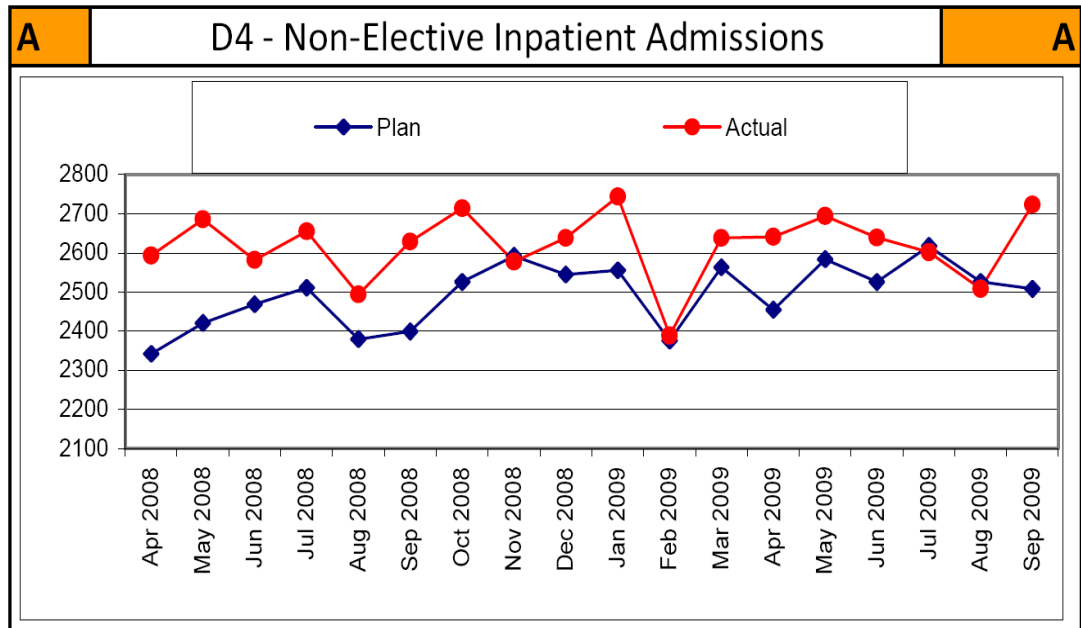


Figure 8.3: CS 1 Emergency (Non-Elective) Inpatient Admissions, Plan vs Actual, April 08 – Sept 09
(extract from CS 1 Trust Board Performance Report, September 2009)

The planned design of the hospital is a separation of the medical and surgical specialities. The wards are designated to treat patients with particular symptoms. For example, within the Department of Medicine there were separate wards with staff who specialise in treating cardiac problems, lung disease, diabetes or elderly patients with multiple illnesses. In the Department of Surgery there were separate wards that specialise in bowel disorders, vascular disease, ear nose and throat, orthopaedics and several others. The planned flow of patients with a particular condition was to the ward that has the specialist staff and equipment to meet their needs. The support services, such as radiology and pathology were shared by all the wards. The planned design of the CS hospitals was a loosely coupled system, which is illustrated by the SFD in Figure 8.4. This illustrates the flow of emergency and elective inpatient admissions into the medical and surgical wards.

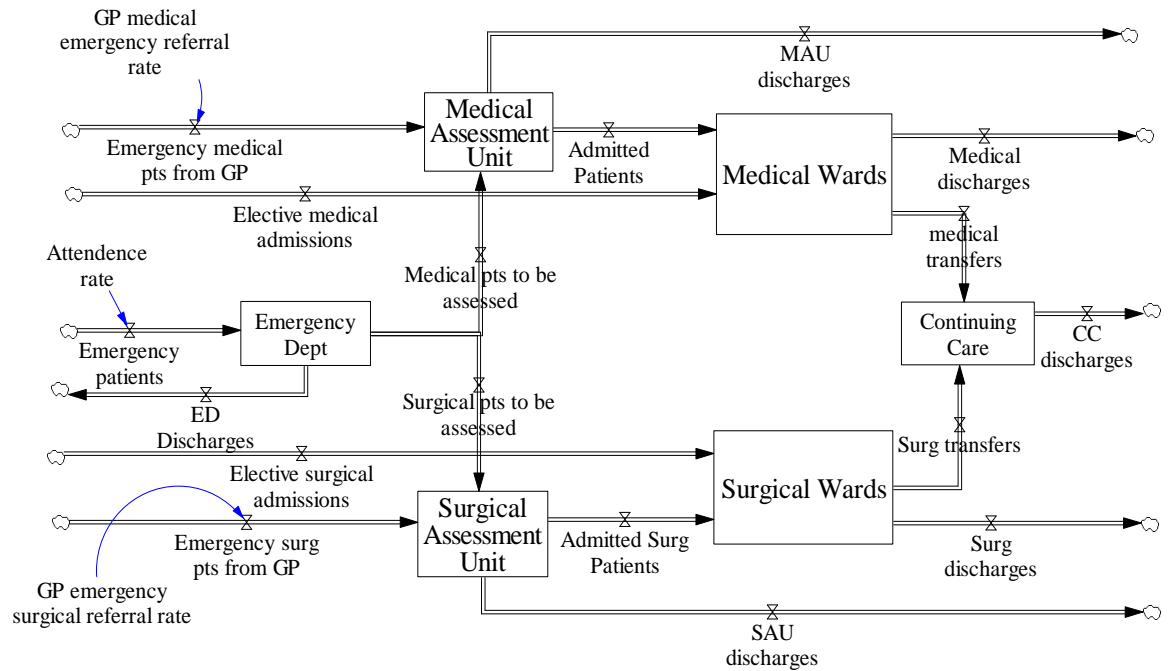


Figure 8.4: The planned stock and flow of surgical and medical patients in CS 1

In this design there are no direct links between the flow of medical and surgical patients within the hospital, apart from in the ED. Therefore, at this level of analysis, the design is largely a loosely coupled system with few feedback loops between the medical and surgical stocks and flows. At a lower level of abstraction there are shared services and staff which do create some feedback loops such as diagnostic and support services. Staff, such as physiotherapists and pharmacists, work across wards. However, for the purpose of this research, the analysis will focus on how the higher level of system design changes when the hospital faces perturbation or a period of continuous stress as described in Chapter 7.

8.3 Factors that affect the planned stock flow design

The ability of the planned design to manage the demand for inpatient services is dependent on the capacity of the stocks (wards) to deal with variation between the rate of flow into and out of the hospital. As noted in Chapter 7 (Figure 7.5), there was a considerable daily variation in the number of elective admissions. There was also a considerable variation in the daily medical admissions and discharges as illustrated by Figure 8.5. The graph shows that for the medical specialities the daily number of admissions can be as high as 85 or as low as 11. There was also a weekly pattern with

lower numbers of admissions and discharges at the weekend and the highest peaks for medical admissions and discharges towards the end of the working week [CS 1 xls 1.7].

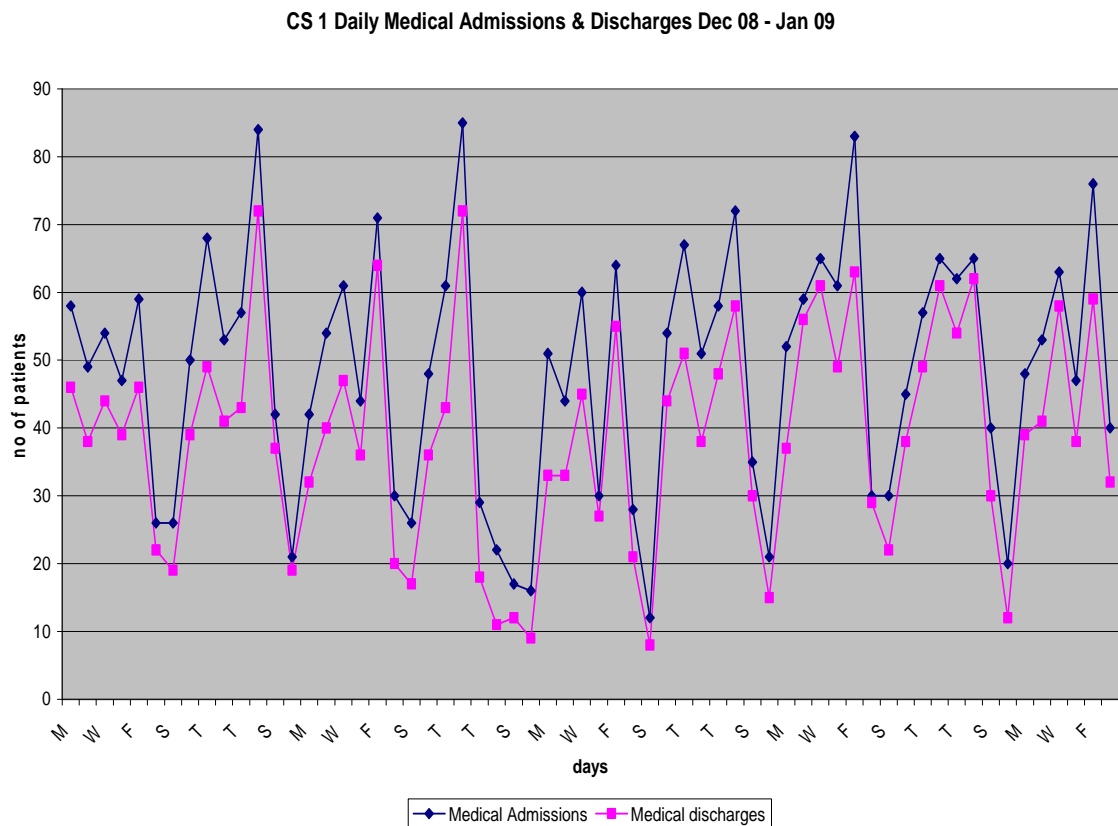


Figure 8.5: CS 1 Daily medical admissions and discharges Dec 08 – Jan 09

At the hospital level there were daily patterns with higher number of admissions than discharges on at the beginning of the week (Sunday) and higher numbers of discharges than admissions on a Friday and Saturday [CS 1 xls 1.11]. This pattern was influenced by the elective admissions which tend to be higher at the beginning of the week and day. This generated an hourly pattern where there were more admissions than discharges early in the day (see Appendix 8.1). Later in the afternoon the discharge rate increases after the consultant ward rounds were completed. Analysis of hourly and daily data for CS 1 using historical data for the period April 06 – March 07 provides the data to illustrate those patterns (Figures 8.6 and 8.7) (Updated data was not available but there was no evidence to suggest any major changes to the patterns.)

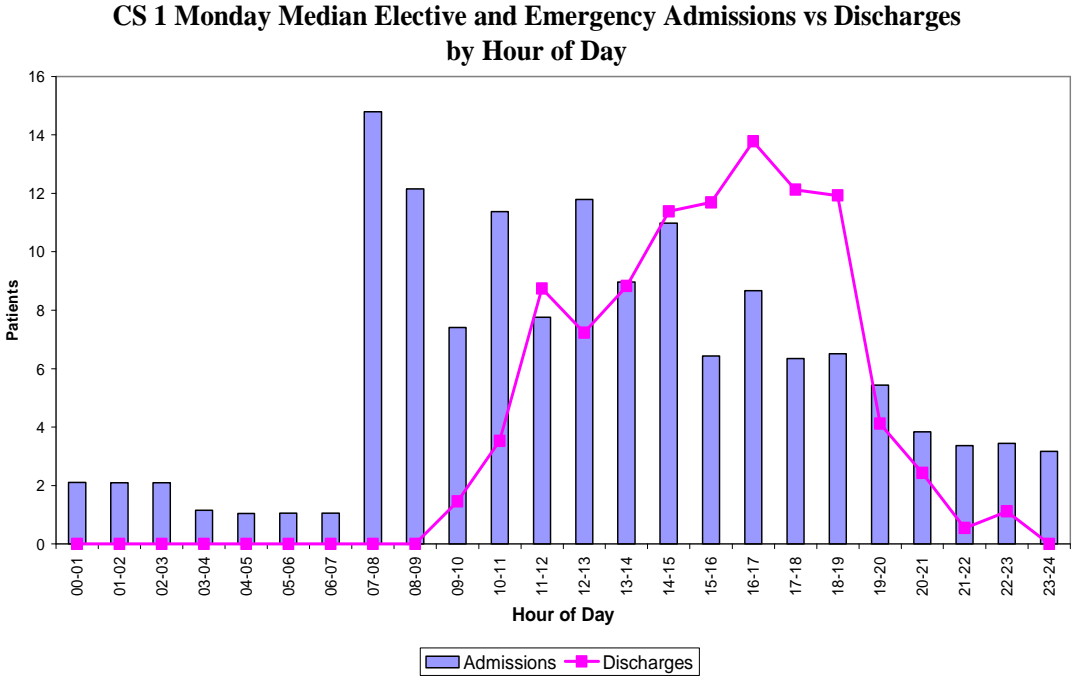


Figure 8.6: CS 1 Monday Median Elective and Emergency Admissions vs Discharges by Hour of the Day 2006-07

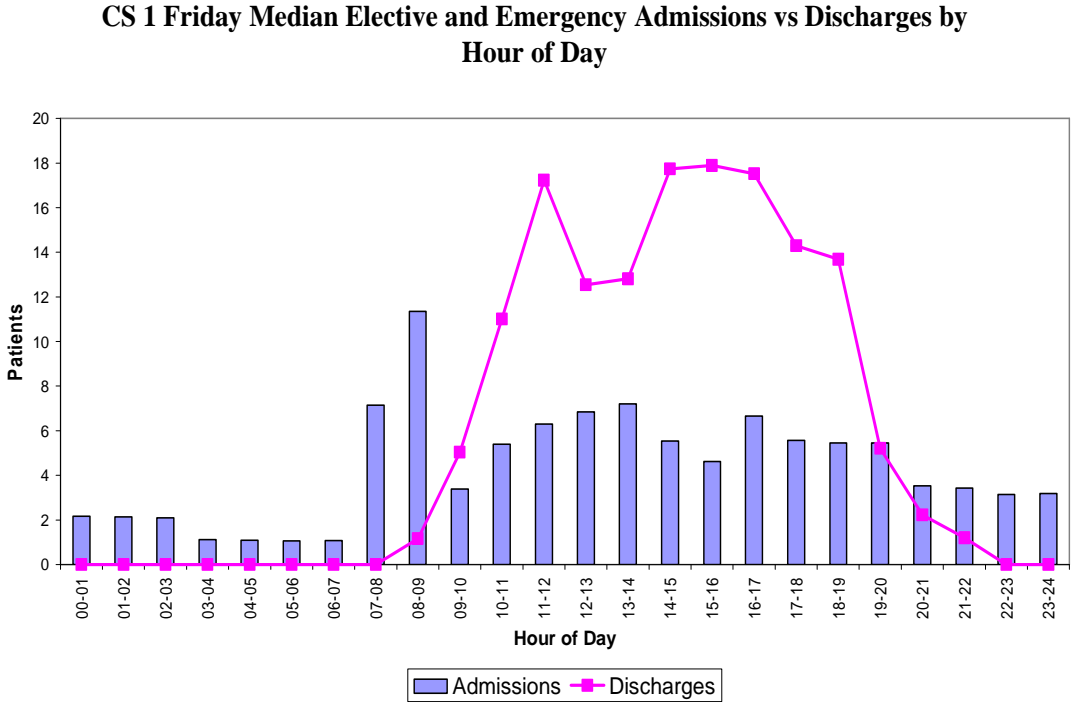


Figure 8.7: CS 1 Friday Median Elective and Emergency Admissions vs Discharges by Hour of the Day 2006-07

The number of patients accumulated in the ‘stocks’ of the hospital depends on the net difference between the rate of flow into (admissions) and the rate of flow out (discharges) of the hospital. The ‘stocks’ absorb the difference between the rates. The

variation between the rate of flow into and out of the system ranges from zero to 200, as can be seen from historical data presented in Figure 8.8.

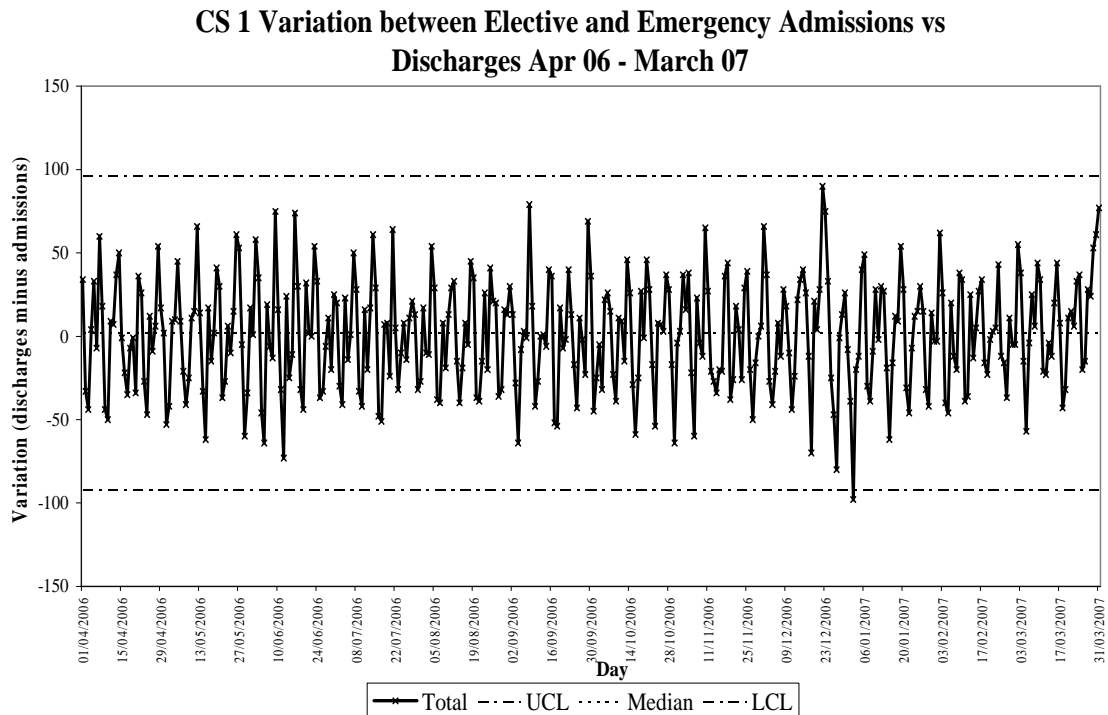


Figure 8.8: CS 1 Variation between Elective and Emergency Admissions vs Discharges April 06 – March 07

For a system to be able to absorb the scale of variation identified, it needs to have a ‘buffer capacity’ of empty beds in the ‘stocks’ (Bagust et al., 1999). When there is disequilibrium between the rate of flow into and out of the system, the hospital has to be able to accommodate the accumulation. As one manager observed:

“So if you think about the patients being in for an average length of stay of may be five or six days on emergency if you have a peak here and another peak six days later that’s not a problem. If you have a peak and then a peak two days later and another peak two days later that first lot of patients are still here. The next lot comes in, the next lot comes in that’s when we hit problems. So it is how close together the peaks are occurring.” (Manager CS 2) [CS 2: RP; DM; Capacity]

Data collected about bed occupancy (occupancy of the ‘stock’) in the NHS is based on a census taken at midnight each day and then averaged over a period of time. Such analysis does not provide the actual bed occupancy of a ward during the day and fails to capture the disequilibrium during the day and week of the in and outflows.

Whilst observing the site management teams it was noticed that most wards were full during the first and middle part of the day with additional patients waiting in corridors or waiting rooms on the ward [FN 2.40-60]. On one particular medical ward the Sister made the following comments:

“...we never go throughout a 24 hour period with a bed staying empty.” (Ward Sister CS 1) [CS 1: RP; DM; Capacity]

“We also have day cases on our ward so every day I have to deal with six to seven day case patients for which I never have any beds for so they end up sat in a corridor for a long period of time (Ward Sister CS 1) [CS 1: RP; DM; Capacity]

An analysis of the data for that Sister’s ward [CS 1 xls 1.10] for the period Nov 07 – Oct 08 shows that the average midnight bed occupancy of the 31 bed ward for the year was 95.3%. An illustrative analysis of the ward data for October 2008 shows that the actual number of patients occupying beds during a twenty four hour period is usually greater than 31 patients (Figure 8.9). This is known as the ‘enhanced occupancy’ rate which is defined as the ‘number of patients per available beds’ during a twenty four hour period (Weissman, 2007). The average midnight occupancy for that month was 97.2%, whilst the average enhanced occupancy was 120.8% (Figure 8.10). This detailed analysis is only for one ward. Observations and interviews indicated that the ward is similar to many of the medical wards in CS 1 and 2 in having an enhanced occupancy rate of over 100% [FN 2.40-60].

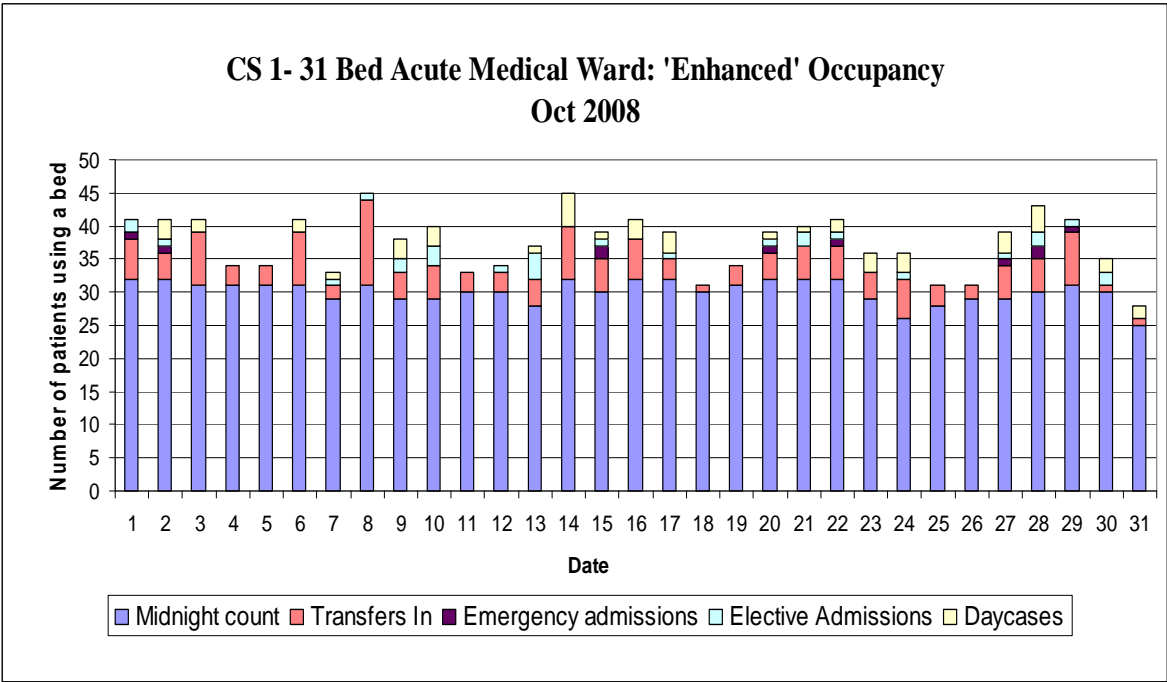


Figure 8.9: CS 1 31 Bed Acute Medical Ward: 'Enhanced' Occupancy (number of patients on ward per day) in October 2008

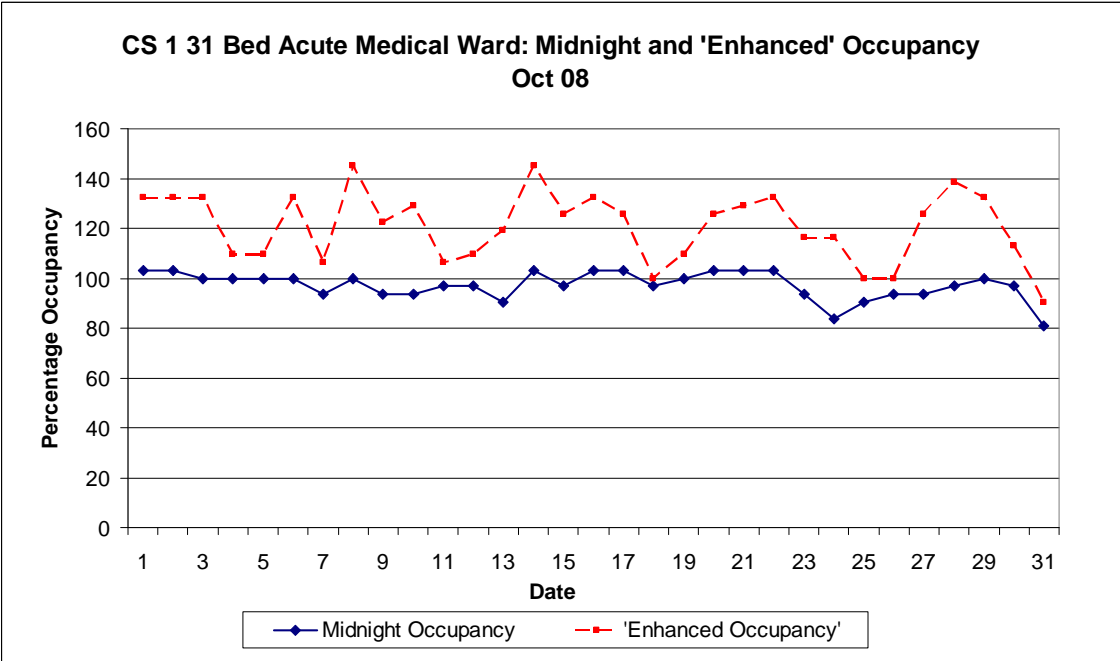


Figure 8.10: CS 1 31 Bed Acute Medical Ward: Midnight and 'Enhanced' Percentage by day - October 2008

In Chapter 7 there is evidence presented about the ED reaching a 'tipping point' when the department became inefficient when it moved close to 100% occupancy. On the

wards there was a similar pattern of the system ‘tipping’ into a situation of ‘inefficiency’.

“The problem I identify with is the lack of flexibility in the system to allow you to actually give optimum care and I mean the problem is the system probably works at its optimum best when you are running at 80-85% capacity if you are running at 100-105% possibly even 110% capacity that creates a huge inefficiency in the system and I think one of the problems is because of financial restraints and this is no criticism of anybody is that the capital investment required to bring in the flexibility has never happened.” (Doctor CS 1) [CS 1: AD; Flexibility]

From this quote the conceptual view is that the system is ‘stretched’ by the need to meet the production (waiting time) targets without spending additional money to fund new ward capacity. The ‘inefficiencies’ will be explored later in terms of the interacting feedback loops that are created by the tight coupling that arises due to the lack of ‘buffer capacity’ in the medical ‘stocks’.

Another factor which reduces the ‘flow’ of patients through the hospital is the design of the medical decision making process. Each patient is under the care of a consultant. The consultant usually has a Registrar, Senior House Officer (SHO) and House Officer (HO) in their team. The key decisions about the patient, particularly related to discharge, are taken by the Consultant or Registrar. In the medical specialities most consultants do two wards rounds a week, albeit there are more frequent Registrar rounds. (EMU has twice daily ward rounds to review recently admitted patients [CS 1: 1.31]) The medical decision making suffers from ‘batching’ which builds in delay to the flow of patients [CS 1: 1.31; CS 2:15]. The situation is made worse when the medical teams have an increased numbers of patients to treat, some of whom may be on surgical wards. As one doctor explained:

“The system does actually become self perpetuating in terms of the log jam in that on EMU we go in every day at eight and we finish our ward round at around half ten/eleven and you can get the jobs done then so you basically create a volume of work investigations normally, take those down to you know the various agencies, CT scanning all these other places and in they are in those departments at ten, half ten whereas if you had the mother of all ward rounds you don’t finish that until half two/three later and actually the working day has gone and by the time you have identified that patient oh gosh they need X, Y, Z scan there’s absolute zero chance you are going to get that on the same day and they’ve you know that’s a big cause of concern and self perpetuating delay in the system.” (Doctor CS 1) [CS 1:AD; Efficiency]

This quote demonstrates that the ward round is the medical decision making process. A key point to note is that the decisions are acted on by the doctors at the end and not during the round. This type of working creates the 'batching' of work thus creating delay in the flow of patients and the wards are not able to discharge patients during the earlier part of the day. Such a situation has a compounding effect adding to the inefficiency.

In summary, there are three issues that put pressure on the planned loose coupled stocks and flows. The first is the large variation in the numbers of patients admitted into and discharged out of the hospital. The second is that the medical wards (stocks) in particular have a high level of occupancy (lack of 'buffer capacity') making it difficult for them to be able to absorb the accumulations of the net difference between the flow rate, without over spilling into non-medical wards. The third is the batching of medical actions due to the ward round method of working. Batching of work creates delays in the flow of patients and creates peaks and troughs of activity for other parts of the hospital system.

In Figure 8.3 the planned design is predominately loosely coupled. However, for the design to be maintained there is a need for stocks to be able to absorb the net difference between the flow rates. Such a 'buffer capacity' is required if the planned loose coupled system is to be sustained in the face of disequilibrium in admission and discharge rates. The evidence from the level of occupancy of the wards was that there is little planned buffer capacity and at times patients are queued to access beds. In the next section the consequences of the very limited buffer capacity for the planned design of the stock flow is explored.

8.4 Tipping from loose to tight coupling

This section examines the impact on the design of the loose coupled system when the medical wards (stocks) do not have the 'buffer capacity' to absorb the net difference in flow rates. There are two changes that occur to the flow of patients when the medical wards become full. These changes are more fully described in Chapters 6 and 7. The first is the diversion of GP medical emergency patients from MAU to ED). The second is the outlying of medical patients onto non-medical wards and into day case and pre-

admission areas. Conceptually, the medical wards are borrowing ‘buffer capacity’ from the surgical wards and day case facilities. The change in flow from medical stocks into surgical stocks increases the coupling between the departments of the hospitals. The hospital moves from being loosely coupled to becoming more tightly coupled with the consequent increases in feedback between the previously separated flows of work. These changes are illustrated in red in Figure 8.11.

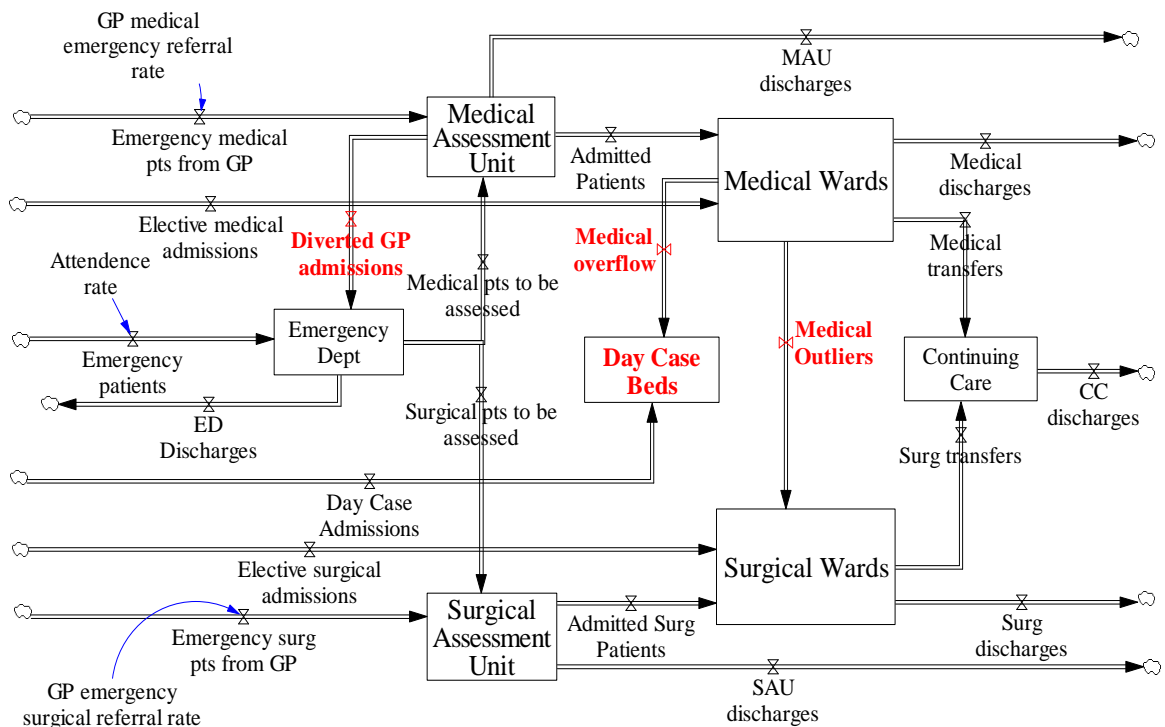



Figure 8.11: Changes (depicted in red) to the planned design of the stocks and flow of surgical and medical patients in CS 1 and 2 when medical wards are full

The implication of the diversion of medical emergency patients to the ED, which is a common occurrence in both CS 1 and 2, is examined first. The diversion generated a feedback loop into the ED which created pressure on the bed capacity, staff workload with potential implications for the safety of patients. The diversion is caused by a net gain of patients due to the disequilibrium of the rates of flow in and out of MAU. It is the lack of empty beds in the medical wards that prevents the transfer of patients out of the MAU. Therefore, the lack of capacity within the medical wards created a feedback loop to the MAU, which caused them to divert GP emergency patients to the ED

The disequilibrium between the rates of flow in and out was too great for the medical wards to absorb. Two reinforcing loops (1 and 2 in Figure 8.12) dominated the hospital

operations at this point. These are illustrated by Figure 8.12 where the symbol  depicts a reinforcing loop. The positive relationship between the variables in these loops means that when the referral rates goes up, so does the occupancy rate. The opposite also applies. Therefore, when the referral rate (rate of flow into the stocks) reduces, then over time, the occupancy of the stocks will decrease as the rate of flow out of the stocks becomes greater than the flow in.

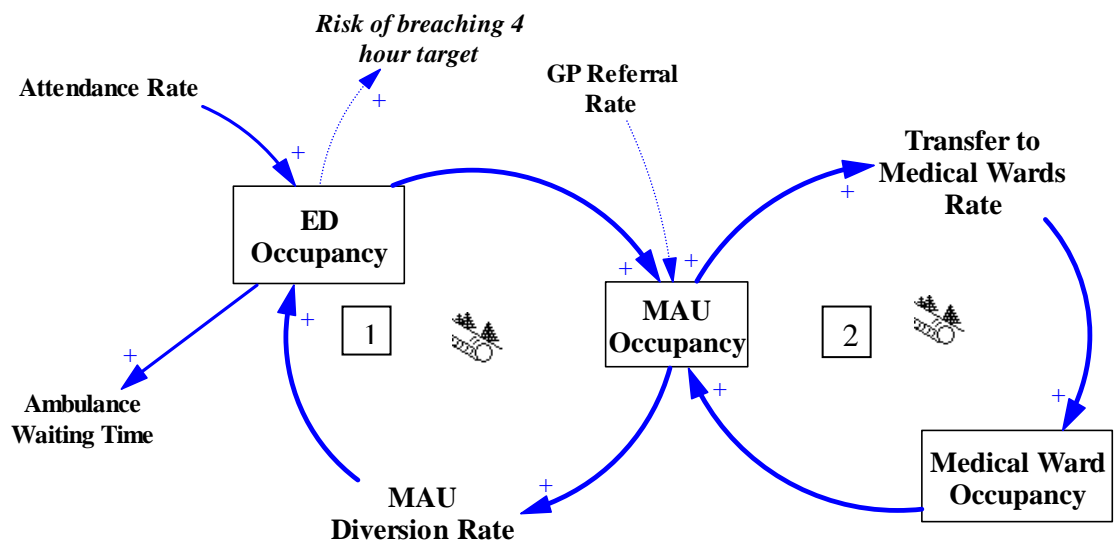


Figure 8.12: CLD of reinforcing loops created in situation where medical ward occupancy is too high to absorb variation in demand. (Stocks in boxes)

The following quote draws out some of the problems for the ED when medical patients are diverted.

“So the medical patients instead of going directly to EMU come here. ...The medical caseload, as I mentioned, have a much higher acuity on average than the emergency lot and that puts huge pressure on the system in terms of bed space so we have to nurse patients in the corridor, we have to monitor patients in the corridor which is particularly vulnerable and we don't have the number of nurses to observe those monitors that we would have per patient should we not have the medical cases. So our pressures primarily relate to the acute medical take.” (ED Doctor CS 1) [CS 1: SWE; Staff workload]

The EDs in CS 1 and 2 are not staffed or equipped to manage GP emergency medical patients in addition to their normal activities. As one manager pointed out:

“If we didn't have the medical and surgical take, so if I could wave a magic wand tomorrow and we didn't have that then I believe that we are staffed

adequately and safely for the patients the ED patients that come through the department.” (Manager CS 1) [CS 1: AD; Staff workload]

The previous year a GP admission medical patient waiting in the ED corridor had died in CS 1. An action following that incident was the purchase of mobile monitoring equipment in an effort to reduce the risk to medically unstable patients of being in the wrong place. It can therefore be argued that the diversion of medical admissions to ED has implications for the safety of the patients.

The medical doctors responsible for the patients diverted to ED were based on the MAU, which in CS 1 is situated a quarter of a mile away. They had to move the initial assessment process from MAU to ED when patients cannot be admitted directly to MAU. The diversion of the flow from MAU to ED produced a feedback loop for the junior doctors. They had to geographically move location and therefore were not available to progress the treatment of the patients on MAU. One doctor described the situation for the initial assessment of patients (clerking) as chaotic:

“On call it just means it’s more chaotic really because the amount of admissions have gone up so we are not really now clerking in MAU any more we are clerking in A&E (ED) and you’ve got everyone sort of going all over the place to try and see the patients and I think as the sort of senior person trying to control the whole situation you have to be really, you really clear about who’s clerking which patient, who’s sick because you can’t review every patient at the moment just because it’s just too busy.” (Doctor CS 2) [CS 2: SWE; Staff workload]

The working conditions for the medical doctors from MAU/EMU in ED were far from ideal:


“...there is a designated amount of doctors to admit people and that way they admit people onto the EMU and they are going to come down onto the ED and do it on there which is what effectively happens. So it’s a stressful and horrible environment when it gets like that because it’s just chocker it’s like a bear pit, ...you can hardly hear yourself above the roar but they still get through the work and yeah, but it will become more difficult.” (Doctor CS 1) [CS 1: SWE; Staff workload]

Conceptually, it can be argued that the unacceptable workload boundary for these key members of staff is in danger of being breached. As they made the initial assessment of the patients there was also a potential movement of the OP towards the boundary of patient safety failure.

The type of feedback loop generated by the diversion of patients was ‘reinforcing’. For example, the patients in the ED were subject to the four hour waiting time rule. The feedback loop generated by the diversion of the medical patients into the ED made the achievement of that target harder, with more patients to be moved through the department within the four hour period (Figure 8.11). The situation created in terms of the overcrowding of the department, staff workload and risk to the four hour target got worse until an action was taken to create a ‘balancing’ loop. A balancing loop is an action that either stops the diversion or allows the rate of flow of patients out of ED to increase. As noted in the two previous chapters, a potential threat to breaching the target boundary attracted considerable management attention and action to resolve the problem. The actions taken to meet the four hour target in these situations can be conceptualised as generating ‘balancing’ feedback loops to increase the flow of patients out of the ED. The actions included the transfer of medical patients to non-medical wards, the use of day case facilities for inpatients and increasing the discharge rate. These are explored in more detail in the next section.

8.5 Increasing the rate of flow out of the medical wards

In CS 1 and 2 the most common action taken to create capacity in the medical wards was to increase the flow of patients out of the wards. This was done either by increasing the rate of discharges out of hospital by encouraging senior doctors to conduct more ward rounds (Figure 8.12, loop 3), or by increasing the transfer rate of medical patients to surgical wards (loop 4). These actions created two ‘balancing’ feedback loops

(depicted by ) that prevent the reinforcing loops (1 and 2) continuing to dominate the hospital system (Meadows, 2008). The balancing loops (3 and 4) are depicted in red in Figure 8.13.

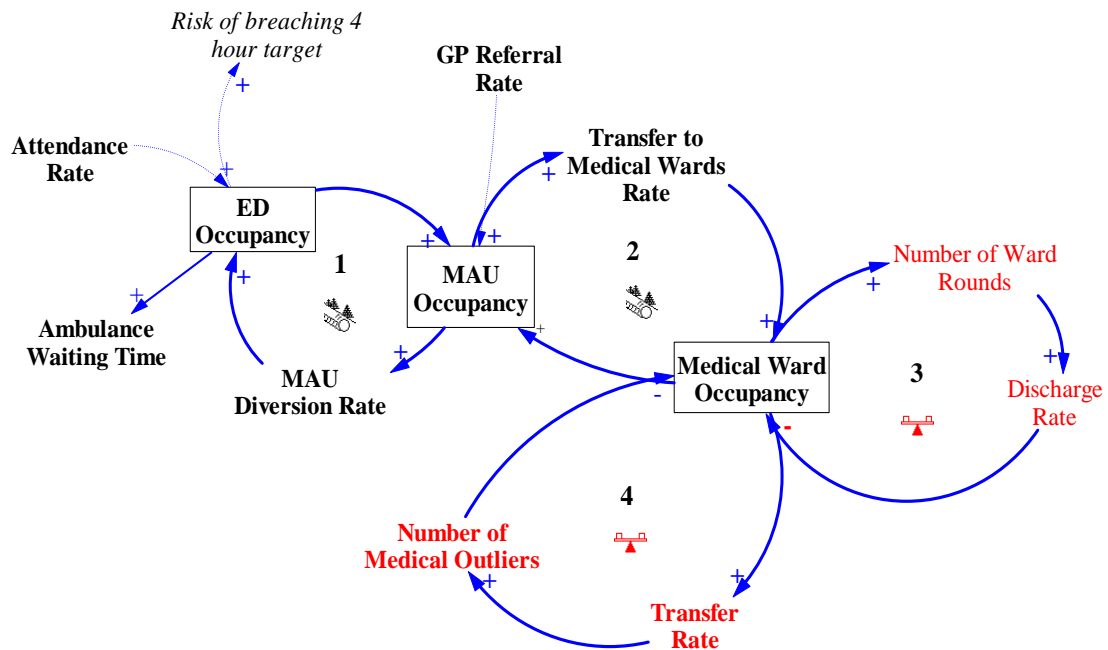


Figure 8.13: CLD with balancing loops of increasing rate of flow out of medical wards

The requirement to instigate the balancing loops was prioritised to the extent that discharges and internal transfers occurred late at night or even in the early hours of the morning. Such a situation is far from ideal for the patient and raised concerns for the staff. However, the staff appeared to recognise that without any buffer capacity to absorb the emergency admissions they had little choice but to move patients at unsuitable times. As one ward Sister commented:

“I never feel comfortable about discharging someone elderly at eleven or twelve at night and you only do that if they are in agreement and the relative’s happy to pick up but, and equally you know you are moving patients off after midnight and they are sound asleep and then you are waking them to say look I am really sorry but we need a bed and we are moving you out to surgery and you want to avoid that if you can and I must say we do it in very small numbers and only when the admission numbers go up rapidly but generally our protocol is it’s one in one out and we identify those patients that are appropriate to move out, there’s nothing worse than you know waking up an elderly confused lady who has probably had sedation and transfer her to a completely different ward and in the morning she wakes up and she doesn’t know where on earth she is. So you know there’s a lot of, there’s a lot of our patient care that just makes you feel you know I know why I’ve got to do it because obviously you know this lady’s fine, she’s got a plan, she’s stable, she’s well someone in ED sitting on a trolley in the corridor is not and needs this bed but equally it just feels uncomfortable what you are doing.” (Wd Sister CS 1) [CS 1:AD; Outliers]

The empty medical beds made available by the actions to increase the rate of flow out are filled by patients from MAU. MAU is then able to accept patients from the ED within the four hour target period. The immediate bed occupancy problem of MAU and

ED is solved by increasing the rate of flow out of the medical wards. However, there are a number of consequences, not all of which were immediately obvious.

Impact of increasing the flow out of medical wards

There are a number of further ‘reinforcing’ feedback loops that were created by the creation of the ‘balancing loops’. These loops impacted upon the medical and surgical services plus the workload of staff. It is suggested that there are also implications for the safety of patients.

In the medical department the primary function of the MAU is for the doctors to assess the patient and decide with them on the treatment plan. With the rapid movement of patients this primary function can be interrupted, which contributes to the inefficiency of the system.

“I have had instances where I have literally been in the middle of the story with the EMU doctor and turned round and they’ve gone to another ward and we have had to wander off and that’s happened over and over and over and over and over again and the inefficiencies of that must be enormous.” (Doctor CS1) [CS 1: AD; Efficiency]

In the wider hospital there is a similar pattern to the characteristics of the ED presented in Section 7.4. When the hospital became overfull the efficiency of the system deteriorated. For example, a key part of organising the discharge of a patient is for the junior doctor to prescribe the tablets to take away (TTAs). The pharmacy uses the prescription to dispense the drugs to the ward prior to the departure of the patient. When ward rounds take longer due to doctors seeing medical outliers, there is a delay in completing the TTA form and getting it to the Pharmacy in time for the patient to be discharged that day.

“When we are in a situation like we are at the moment and the whole of medicine becomes so inefficient with lots of patients all over the hospital junior doctors go all round their outliers and their in-patients with their consultants and they are delayed in getting discharges done and complete and of course we, our TTAs are on the back of the discharge summary. So you know by the time the juniors get back at 3 o’clock in the afternoon to do a discharge summary.” (Manager CS 2) [CS 2: AD; Efficiency]

Increasing the transfers out of medical wards may actually decrease the rate of discharges from the hospital. Using surgical beds for medical outliers does however,

provide a short-term solution to the capacity issues of the medical wards to accept patients from MAU. Figures 8.14 and 8.15 demonstrate that outlying patients is almost a daily occurrence. It is suggested that staff have normalised to the practice of borrowing capacity. The data from CS 1 shows that it was also an increasing procedure.

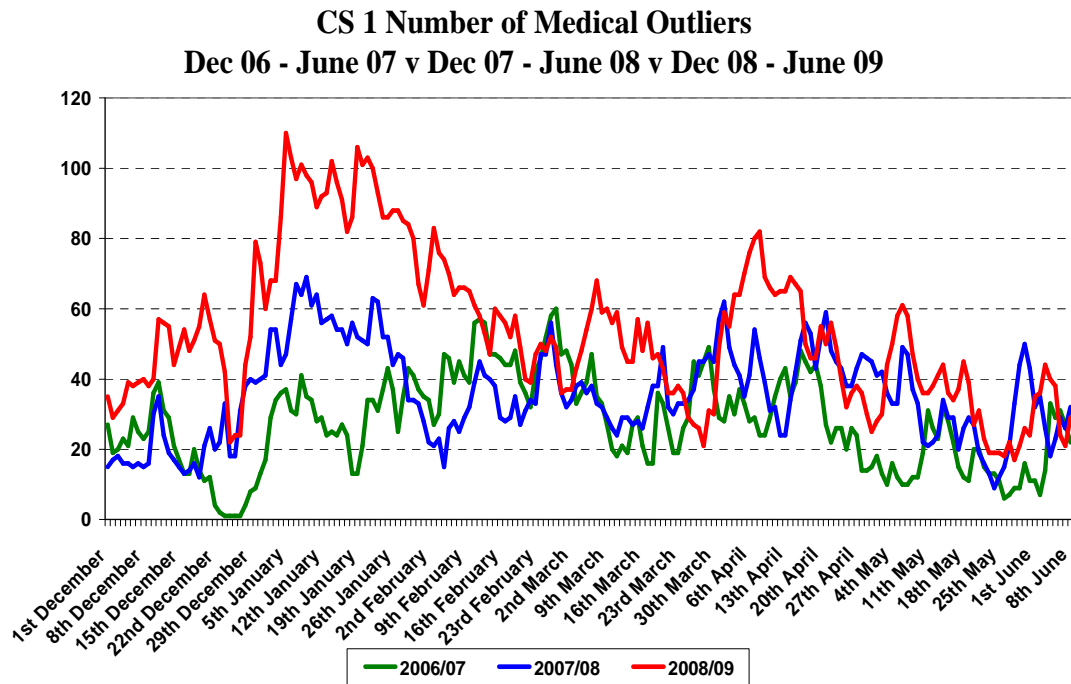


Figure 8.14: CS 1 Number of Medical Outliers Dec – June 2006-2009

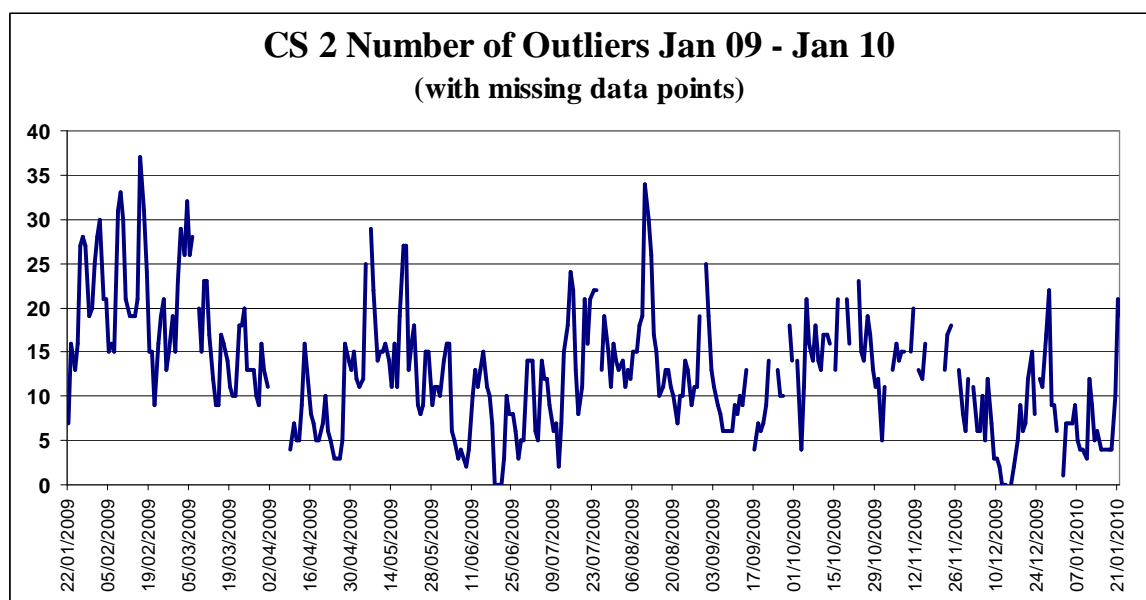


Figure 8.15: CS 2 Number of Outliers Jan 09 – Jan 10 (with missing data points)

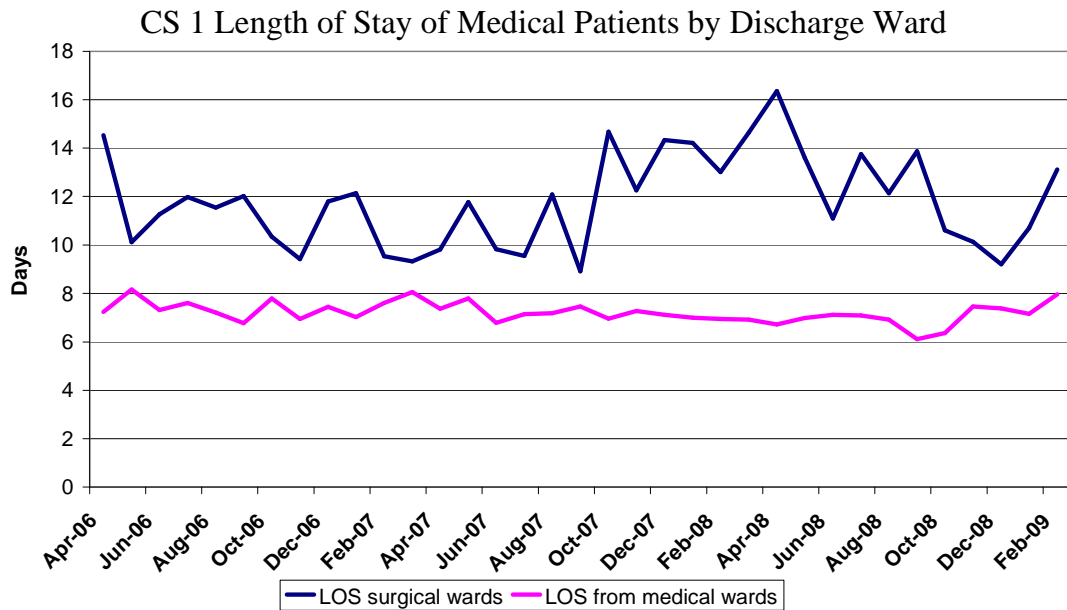


Figure 8.16: CS 1 Length of Stay of Medical Patients by Discharge Ward

Medical patients who are outlied in surgical wards have a considerably longer length of stay than those who are not outlied (Figure 8.16) [CS 1: xls 1.9]. There are a number of potential reasons for that difference. One explanation is the selection of patients to be outlied can be biased towards those with a normally longer length of stay. Equally, it is recognised that the patients often change medical team when they move wards. When the handover of the patient occurs then it can take time for the new consultant team to get to know the patient and therefore manage their treatment efficiently. Therefore, the transfer of patients introduces the potential for delay and failures in communication. The change of medical team even from MAU to the medical wards was not always a seamless or timely process, raising concerns about the potential safety of patients:

“What happens is I think the nurses get a handover the patient turns up on the ward and then the ward clerks allocate them a team but essentially they are not allocated a team before five or they can for four thirty they won’t actually have a team so they won’t be seen any of the juniors on the ward. So you have to hope that they are sorted by the MAU and what you do sometimes find is that not everything is done on MAU and they are then seen the next day by us and as long as they are not sick that’s fine it’s just chasing things up but I suppose.” (Doctor CS 2) [CS 2: AD; Continuity]

There was also the impact of medical outliers on staff workload. The physical bed capacity needed to absorb the net difference in rates of flow was borrowed by the

medical wards from surgery. However, the ‘buffer capacity’ generated by this action also increased the workload of staff, particularly the doctors, senior nurses and managers. The key balancing loop of increasing the transfer and discharge rate from the medical wards often relied on asking staff to work harder, both in undertaking additional ward rounds, and then taking all the required actions to process the patient. The use of staff in this way can be conceptualised as borrowing buffer capacity from the staff by moving the unacceptable workload marginal zone boundary outwards.

In CS 3 there is evidence from the inquiry reports that it was the shortage of staff in key stock areas (ED, MAU, Medical Wards) that created the conditions for breaches in the patient safety boundary (Francis QC, 2010a). Conceptually, the exhaustion of the buffer capacity associated with the unacceptable workload boundary allowed the OP to breach the patient safety boundary in that hospital as the staff did not have the capacity to maintain safe practice.

Increasing the rate of flow of patients out the medical wards through medical outliers did create a number of consequences. The creation of the ‘balancing’ loops (3 and 4), generated a number of reinforcing loops, illustrated in Figure 8.17.

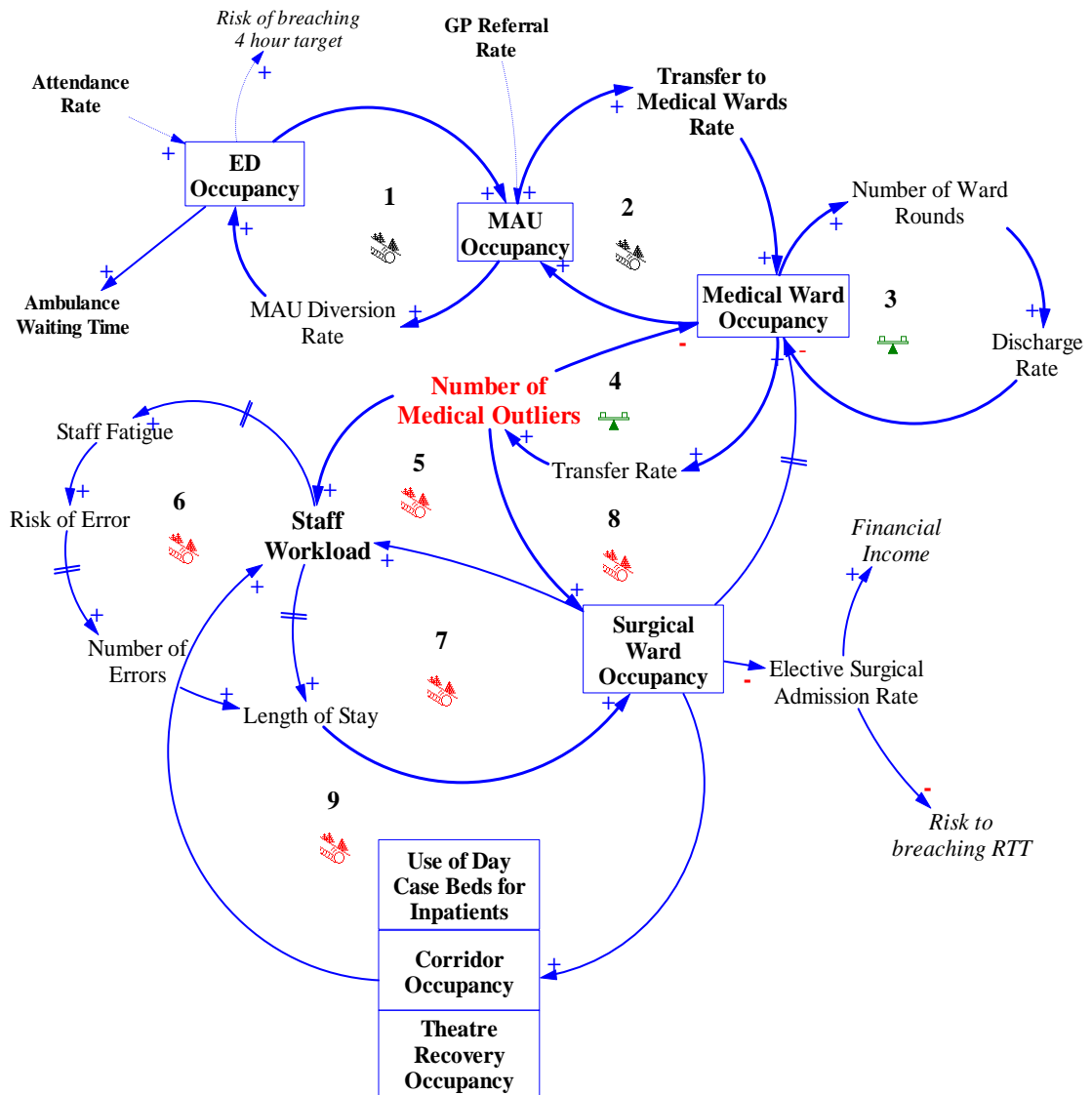







Figure 8.17: Reinforcing feedback loops generated by ‘medical outliers’

The first of the ‘reinforcing loops is created due to the increased staff workload for doctors in particular, having their patients on more than one ward (loop 5). Higher workload for staff is associated with increased fatigue which increases the risk of error (loop 6) (Hall et al., 2004, Rogers et al., 2004, Scott et al., 2006). When errors do occur the resultant harm usually means an extended length of stay for the patient (Sari et al., 2007b). With a higher number of medical patients to look after spread around the hospital, the ability to see, treat and discharge patients slows down. After a delay this situation leads to an increase in the length of stay (see Figure 8.16), which impacts on the occupancy of the medical and surgical wards (loop 7 and 8). This situation has potential consequences for the elective surgical admission rate and associated RTT waiting time target and financial income. The threat to the RTT and financial income

was responded to by admitting more patients than there were beds for and using corridors and theatre recovery as temporary holding areas for patients [FN 2.40-60]. The safety considerations of these actions did not appear to take a high priority.

Some of the other loops have delayed outcomes and are therefore not always noticed by decision makers or regarded as linked to the number of medical outliers (Booth Sweeney and Sterman, 2007). The feedback loops 5 - 9 in Figure 8.15 are interrelated. The explanation of the loops 5 - 9 is as follows:

Loop Number	Explanation of relationships
5	 Reinforcing loop 5: when the 'number of medical outliers' goes up the 'staff workload' goes up. As 'staff workload' goes up their ability to treat patients efficiently decreases and there is a delayed increase in the 'length of stay' (LOS) for patients. When the LOS increases the 'surgical ward occupancy' increases. When the 'surgical ward occupancy' increases there is a delayed impact on the 'medical ward occupancy' as fewer surgical empty beds are available to transfer patients into. However, as the 'medical ward occupancy rises' so does the 'transfer rate' increase. The surgical wards are then forced to create additional stock capacity (see loops 8 and 9)
6	 Reinforcing loop 6: When the 'staff workload' increases the 'staff fatigue' increases which raises the 'risk of error' occurring. With an increased risk, there is the likelihood of an increased 'number of errors'. Errors create harm for patients who then stay longer in hospital, increasing the 'LOS'. Loop 6 then feeds into loops 7 and 8
7	 Reinforcing loop 7: When the 'LOS' rises the 'surgical ward occupancy' rises which in turn increases the 'staff workload'.
8	 Reinforcing loop 8: As the 'number of medical outliers' increase then the 'surgical ward occupancy' rises. When the 'surgical ward occupancy' increases there is a delayed impact on the 'medical ward occupancy' as fewer surgical empty beds are available to transfer patients into. This situation generates loop 9.
9	 Reinforcing loop 9: When the 'surgical ward occupancy' rises then, in a similar way to the medical wards, the rate of discharge increases (not shown) or additional inpatient capacity has to be created. This is depicted as the 'use day case beds for inpatients', 'corridor occupancy' and 'theatre recovery occupancy'. When these additional 'stocks' are used it increases the 'staff workload'.

	When the 'surgical ward occupancy' rises due to the transfer in of medical patients, then there is the potential for the 'elective surgical admission rate' to decline, due to the shortage of beds. When the elective admissions decline the financial income for the hospital reduces and the risk of breaching the Referral to Treatment Time (RTT) increases.
--	---

Table 8.1: Explanation of loops 5 – 9 in Figure 8.16

Conceptually, it can be argued that the need to admit emergency medical patients and elective surgical patients to keep the OP away from the financial and target failure boundaries placed considerable pressure on decision makers to stretch the physical and staff capacity. Data from the staff indicate that it was the patient safety failure gradient that motivated them to prioritise admitting emergency patients within the 4 hour time period. However, the implications of this prioritisation in terms of the tight coupling and the nature of the feedback loops generated did not appear to be well understood.

The 'balancing' loop of increasing the rate of discharges / transfers

The 'balancing' loop of generating medical outliers created a series of dynamic interactions that allow reinforcing loops to dominate the hospital system. At a system level, the balancing loop that can bring the system back into a stable situation is to create a net reduction in the flow rate into the hospital. The rate of discharges has to be greater than the rate of admissions.

The action to prevent the reinforcing loops domination is the action to create the 'balancing' loop of increasing the discharge rate. This action can reduce the LOS of patients (increase the rate of flow out of the system) and consequently reduce the ward occupancy. This action stops the need for medicine to borrow 'buffer capacity' from surgery. There are a number of feedback loops that influence the 'length of stay' in the hospital system as shown in Figure 8.17. Figure 8.18 illustrates in more detail some of the feedback loops generated by the medical wards (stock) borrowing 'buffer capacity' from other parts of the system. The CLD illustrates that there are potentially four reinforcing feedback loops that will dominate the system without a sufficiently powerful balancing loop (loop 5).

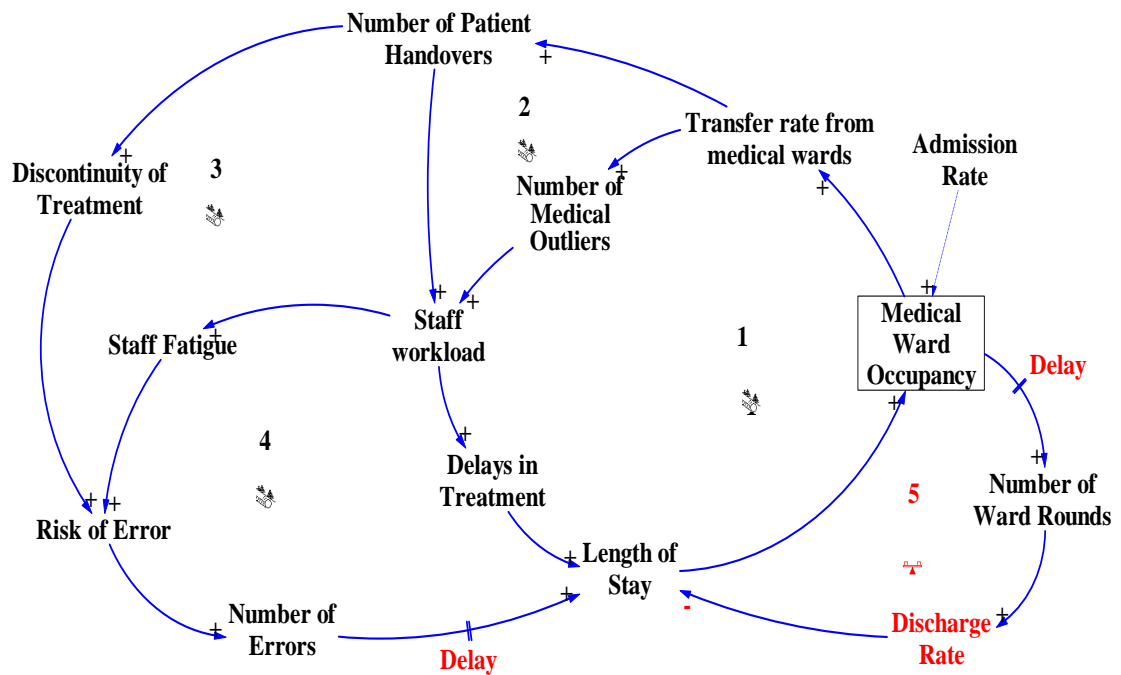


Figure 8.18: Feedback loops generated by transfer of medical patients to surgical wards

There was considerable pressure for the balancing loop to maintain the overall stock level with the capacity to manage the inflow of patients. At very busy times consultant medical staff increased the number of ward rounds and some seek to discharge patients slightly earlier than they would normally do so. Some doctors put safeguards, such as an early outpatient or diagnostic appointment, in place.

“I think there’s a heightened awareness that we need to be getting patients out of hospital quickly and efficiently and so are often sending people out of hospital with and are more likely to build in an out patient follow up as a safety net because we are sending people home perhaps a bit earlier than we would have done.” (Doctor CS 1) [CS 1: AD; Flow]

In both CS 1 and 2 there were a number of patients who are medically fit for discharge. However, factors beyond the control of the hospital, such as the lack availability of continuing care in community hospitals or nursing home beds, means that the patients remained in the acute hospital ‘blocking’ beds. Staff working in social services or community based NHS services did not easily perceive the feedback loop that the delay in their actions generate on the stock and flows in the acute hospital. Conceptually, the OP of the hospital system is therefore influenced by the management of stocks and flows in continuing care. Continuing care systems tend to have less production and more financial control focused priorities.

In the next section the conceptual aspects related to the SWE (v3) derived from this analysis are identified.

8.6 Conceptual analysis



In this section the dimensions of the ‘structure’ and ‘feedback’ constructs of the SWE (v3) model are identified. The implications on the movement and location of the OP within the model are also explored.

8.6.1 ‘Structure’

The first dimension of the ‘structure’ construct is ‘coupling’. As noted in Section 8.2 the planned design of a hospital was for a separation between the medical and surgical flow and stocks of patients. The planned design of the hospital can be described as loosely coupled. However, when the medical wards (stocks) were full, the flow of patients was diverted to the ED or into surgical wards. The planned design of separation was changed to one of the flows being linked together. Using SD terms, when the flows and stocks of medical and surgical patients become connected, the parts of the system move from ‘loose’ to ‘tightly’ coupled.

The second dimension of the ‘structure’ construct is the ‘buffer capacity’. This dimension depicts the ability of the ‘stocks’ to absorb the net differences between the flows into and out of the system. As noted in Chapter 2, the stocks are the accumulation of the net differences between the rates of flows in and out. In Section 8.3 it is shown that there was considerable variation in the rate of flow into and out of the hospitals studied. It was also noted that the medical wards (stocks) in particular had a consistently high occupancy level. This means that the medical wards were often not able to absorb the variation in flow rates in and out without having to borrow ‘buffer capacity’ from other parts of the hospital.

8.6.2 ‘Feedback’

As noted in Chapter 2, there are two types of feedback loops; therefore the dimension are either reinforcing (positive)  or balancing (negative) . There are implications for the movement and location of the OP from the type of feedback loop that dominates the system.

The OP in the SWE (v3) depicts the behaviour of the system in relationship to the boundaries. The connection between the structure and behaviour of a system is understood by linking the feedback process with the stocks and flows (Sterman, 2000). A system is in ‘dynamic equilibrium’ when all the flows in and out are in balance. However, few complex dynamic systems, such as hospitals, have such equilibrium due to the variation in flows. The ‘stocks’ accumulate the net rate of change in flows. Non-linear feedback loops means that there can be a shift in the type of feedback loop that dominates the system.

When a system has a ‘balancing loop’ dominant then the rate of flow out is greater than the rate of flow in. The system is in a ‘stable’ condition. When this is applied to the SWE (v3) model it can be argued that when the ‘balancing loop’ is dominant, then the ‘movement’ of the OP is ‘stable’. However, the system can change into a situation where the ‘reinforcing’ loop dominates. This occurs when the rate of flow in is greater than the rate of flow out. The ‘movement’ of the OP can be conceptualised as being ‘unstable’ when reinforcing loops dominate the system. The dimensions are summarised with examples in Table 8.2:

Construct	Dimension	Examples
Structure	Coupling	Planned design was loose coupling of medical and surgical stocks and flows. Medical outliers created tight coupling between otherwise separate stocks and flows.
Structure	Buffer Capacity	Buffer capacity for medical patients was borrowed by the use of surgical beds. The staff worked harder to create and maintain the borrowed buffer capacity
Feedback	Reinforcing or Balancing	Diversion of emergency medical admissions to ED and the use of surgical beds for medical patients generated a series of reinforcing feedback loops. Increasing the rate of discharges was the only system level balancing loop available to bring hospital back into a stable position.

Table 8.2: Dimensions of the ‘structure’ and ‘feedback’ constructs of the SWE (v3)

8.6.3 Implications on the OP

Where the rate of flow into a system is greater than the flow out, the stocks absorb the net difference until they reach capacity. As noted above, the CS hospitals became overfull until such time as rate of flow out became greater than the flow in. It was the ability of the staff to flex their workload that provided much of the buffer capacity to sustain the hospital working at above the physical bed capacity. There was often a delay in the 'balancing' loop becoming dominant, which created the situation where the hospital operated, for a period of time, at over capacity. This situation lasted for a matter of a few hours or days. Due to the delay, the 'balancing' loop can often over correct the state of the stocks. This results in the system oscillating between a 'stable' and 'unstable' situation.

When a hospital system experiences perturbation or continuous stress then it can reach a 'tipping point'. It is at that point that it switches from a dominant 'balancing' loop to a 'reinforcing' loop, which can drive exponential growth (Sterman, 2001). Conceptually, the 'tipping point' can be depicted as the 'movement' of the OP switching from 'stable' to 'unstable'. When the OP is 'unstable' it can move quickly and breach any of the envelope boundaries.

Conceptually, this 'tipping point' can apply to parts within the hospital system. As noted in Chapter 6, one of the doctors in CS 2 described the ED reaching a 'tipping point' when it became overfull and consequently inefficient [CS 2.9]. The 'stability' of the ED OP was recovered in those types of situations when the 'balancing loop' of increasing the rate of flow out of the department reached the point where the number of patients coming in is less than those leaving.

It is therefore argued that by assessing the stock, flows and feedback loops, valuable insights are gained as to the dynamic behaviour of a complex socio-technical system such as a hospital. By using SD as part of the SWE (v3) model it can take account of how the system behaves when the flow of work changes. Conceptually, insights are gained as to what happens to stocks, flows and feedback when decision makers take actions that keep the OP away from certain boundaries. The application of SD in this study is limited to a high level analysis and has not sought to develop a simulation model of the hospital due to the large number of variables involved.

8.7 Summary

This chapter examines the ‘structure’ and ‘feedback’ constructs of the SWE (v3) model by exploring the stocks, flows and feedbacks found in CS 1 and 2. The planned design of the separation of the stocks and flows for medicine and surgery was compromised when the department of medicine stocks overflowed into surgery, creating tight coupling of the hospital. The overflow is the result of the medical wards not being able to absorb the net difference in the rates of flow in and out of their part of the hospital. There are a number of feedback loops that became active when the system becomes tightly coupled. Individual departments and the whole hospital reached a ‘tipping point’ when there was a switch from the ‘balancing’ to ‘reinforcing’ loop dominance.

The dimensions of the ‘structure’ and ‘feedback’ constructs are identified. Two important conclusions about the characteristics of the hospitals can be inferred from the analysis. The first conclusion is the key role of ‘buffer capacity’. Where there is sufficient ‘buffer capacity’ to absorb the net difference in the rates of flow, then a system can maintain the planned design. The second conclusion is that the system can reach a ‘tipping point’ due to the dynamics of the flow rates creating a switch in the type of loop dominating the system. When ‘reinforcing’ loops dominate then the system is unstable and the OP can make large movements.

In the case studies there was evidence of a chronic shortage of ‘buffer capacity’ in the medical ward stocks. The result is that ‘buffer capacity’ was frequently borrowed from the surgical wards and from the ability of staff to work harder. What was not perceived clearly by decision makers was the potential ‘drift to danger’ created by constantly flexing the unacceptable workload boundary to absorb the difference in flow rates and generate ‘balancing’ actions.

In the next chapter the results and analysis presented in Chapters 6, 7 and 8 are discussed in the light of the extant literature and proposals are made about how to improve patient safety from a system resilience perspective.

Chapter 9 – Discussion

9.1 Introduction

The pragmatic critical realist position argues that it is legitimate to access knowledge through observing the effect of a directly unobservable phenomenon. The ‘system’ within the case studies is not directly observable. However, the actions of the hospital staff and workload were observed, analysed and are presented in Chapters 6-8. The constructs of the SWE (v3) model are identified and operationalised to provide insights into the characteristics of the system and their influence on the risk to patient safety. This Chapter makes the link back to the research objective and key points in the literature to set the context for the discussion about the characteristics of the system.

The research objective is to explore, in NHS hospitals, how a systems approach can inform the development of patient safety theory. In Chapter 3 a system resilience model (SWE v1) is derived from the literature (Rasmussen, 1997). The limitations are noted and Chapter 5 sets out the development and contextualisation of the model to provide the vehicle to explore patient safety in the context of the NHS in England. Chapter 6 to 8 present the analysis of the data from the case studies using the SWE v3 model.

A major part of this research has been in the development and use of a system resilience model to gain an insight into system behaviours and how they might influence patient safety. Section 9.2 discusses the development and insights gained from the SWE v3 model. In Section 9.3 the construct sets and their interactions are reviewed and propositions made about the emergent system behaviours. It is argued that some of those emergent behaviours are problematic for the safety of patients. Five patterns of system behaviour are proposed, four of which are considered to increase the risk to patient safety. Ideas as to how to improve the system resilience to reduce the risk are made. Finally, the contribution to knowledge is summarised.

9.2 The development of the Safe Working Envelope model

The development of the model has to be set within the context of the research objective which is to explore how a systems approach can be used to provide an insight into patient safety in NHS hospitals. There are a number of points from the literature which

have been incorporated into the development of the SWE v3 which are summarised. Using the constructs and dimensions of the SWE v3 model to conduct the case studies has provided a number of insights into some of the characteristics of the hospitals studied.

9.2.1 Key points from the literature

The extant system models found in the patient safety literature do not apply concepts from systems thinking to consider the implications of the dynamic interactions that occur in healthcare organisations. For example, the frameworks suggested by Vincent et al, (1998) and Donabedian (1966) provide long lists of system factors to be taken into account when considering patient safety but do not explore the potential interactions. The popular Swiss Cheese model (Reason, 1997) takes a view of systems which is based on avoiding linear component failures. Therefore, the extant literature lacks explanatory power about patient safety in complex hospital systems.

Complex socio technical systems have particular characteristics (Cilliers, 1998), which have to be taken into account when developing a theoretical framework to improve the understanding about how the ‘system’ influences patient safety. It is argued that insights from systems thinking, resilience and accident theory can provide a different theoretical framework through which to analyse safety failure. A systemic approach, rather than the linear component paradigm, provides greater explanatory power about how the system influences patient safety (Dekker, 2011). There are a number of points that arise from the synthesis of the literature about the characteristics of ‘systems’ that assist in developing the understanding of patient safety.

The first is that healthcare organisations, such as hospitals, are complex socio technical systems with open and fuzzy boundaries. There are multiple interactions between the social and technical aspects. Therefore, insights about patient safety have to commence with the knowledge that safety is an emergent property of a complex set of interactions. Systems can be structured in ways that the resultant interactions produce problematic behaviour (Meadows, 2008) The development of system resilience models, such as Rasmussen’s (1997) SWE, provides a means to gain insights into the behaviour of complex socio technical systems.

The second point is that a rise in the numbers of interactions between the parts increases the complexity of the system (Perrow, 1984, Cilliers, 1998, Dekker, 2011). SD theory

provides a means to gain insights into the interactions of the parts and the presence of feedback loops (Sterman, 2001, Meadows, 2008). Increased complexity requires a corresponding rise in the adaptive capacity of those working in the system to avoid things going wrong (Woods et al., 2007, Woods et al., 2010, Ashby, 1961, Dekker, 2011). As noted, the increased complexity also makes it more difficult for decision makers to appreciate the dynamic relationships due to the delays in feedback (Booth Sweeney and Sterman, 2007).

The third point is that resilience theory provides a way to consider safety at a system rather than component or task level (Hollnagel et al., 2006). The ability of a system to anticipate, adapt, absorb and recover from a disturbance or continuous significant stress is taken into account. The SWE model includes the adaptive capacity of the system, depicted by the marginal zones of the envelope. Conceptually, the adaptive capacity is deployed through compensating actions in an attempt to hold the OP in a stable position inside the performance envelope.

The fourth point is that the extant accident theory literature provides a number of concepts that help to explain why accidents happen. It is argued, that those concepts can be unified by the use of the SWE model from the ‘resilience engineering’ literature (Rasmussen, 1997, Cook and Rasmussen, 2005). As identified the SWE v3 provides a unifying model to incorporate the concepts derived from the accident theory literature. These are summarised in Table 9.1.

Concept	Principle Literature	Model construct	Examples from case studies
Production verses safety	(Reason, 1997, Flin et al., 2008, Dekker, 2005)	Boundaries, gradients, location of OP	Not cancelling elective admissions during sickness virus
Blunt / front end	(Reason, 1997)	Pressure from gradients; response of staff	EWTD on junior doctors; prepared to work longer to gain training opportunities
Latent or hidden conditions	(Reason, 1997)	Competing pressures on the OP	Combination of waiting time targets impact on patient in ED
Safety as a dynamic non-	(Weick, 1987)	Gradients and compensating actions	Individual actions of clinicians to keep

event			patients safe
Coupling & feedback	(Perrow, 1984, Diehl and Sterman, 1995)	Stocks, flows & feedback loops	Medical outliers increasing length of stay
Redundancy / buffer capacity	(Miller and Xiao, 2007)	Location of the marginal zone boundary	97.2% achievement against 98% ED target
Normalisation	(Vaughan, 1999)	Location of the OP	Patient falls
Practical drift	(Snook, 2000)	Movement of the marginal zone boundary	Using gynaecology ward to outlie medical patients
Trade-offs	(Hollnagel, 2009a)	Making visible the boundaries and OP location	ED admissions traded off with medical outliers

Table 9.1: Concepts from accident theory and their relationship to the SWE v3 model constructs.

9.2.2 Insights about the constructs and dimensions of the SWE v3 model

The use of the SWE (v1) model depicts a complex socio-technical system operating within the constraints of the failure boundaries. However, the envelope model is limited and is therefore extended by the use of SD to provide an insight into the dynamics found inside the envelope that influence the operating point of the system. The principle underlying systems thinking is that the interactions between the parts of a system contribute to the dynamic of the whole (Forrester, 1961). In the extant literature, SD is used to study the operational management of hospitals but makes very limited if any links to the implications for the safety of patients (Lane et al., 2000, Lane and Husemann, 2008b, Winch and Derrick, 2006). This thesis develops and contextualises the conceptual model to allow the connection to be made between the system dynamics of operational management, the resilience of the hospitals to perturbation or continuous stress, and the resultant implications for patient safety.

The SWE model is conceptualised as a ‘system resilience’ model (Rasmussen, 1997, Woods et al., 2009, Cook and Rasmussen, 2005). It is argued conceptually, that a resilient system is one that can keep the OP within the failure boundaries of the safe

working envelope. The extant literature assumes that the OP is normally located close to the marginal zone boundary for patient safety failure (Cook and Rasmussen, 2005, Miller and Xiao, 2007, Amalberti et al., 2006, Reason, 2008). There are a number of reasons for this assumption.

- The costs associated with having the OP well away from the safety marginal zone are high in lost production and money (Reason, 1997)
- The pressure exerted from the gradients towards efficiency (finance) and least effort (unacceptable workload) are stronger than the safety gradient (Rasmussen, 1997)
- The safety boundary is the least well defined in comparison to the other boundaries (Rasmussen, 1997)
- The work required to maintain safety is often unseen by senior decision makers who stretch the system to ever greater efficiency (Woods, 2006)

The findings from the empirical case studies agree with the literature on these points. This research also suggests that pressure exerted, from the target boundary gradient in particular, creates system dynamics that normally locate the OP inside the marginal zone of the patient safety boundary or move the marginal zone boundary outwards.

SD can help to provide insights into the system characteristics that contribute towards the OP in CS 1 & 2 being normally inside the marginal zone. For example, a finding from the SD analysis of the case study data is that the MAU / EMU and medical wards do not have sufficient 'buffer capacity' to adapt to the variation in demand. Using SD terms; the 'stocks' do not have the 'buffer capacity' to accumulate the net difference between the in and out 'flow' rate. The lack of 'buffer capacity' makes the system vulnerable to having to borrow capacity. This finding is considered further in Section 9.2.3. Medical patients flowing into surgical beds creates conditions of overcrowding that moves the location of the OP to be close to if not breaching the patient safety boundary (Fatovich et al., 2005, Sprivulis et al., 2006, Trzeciak and Rivers, 2003, Wears et al., 2008, Weissman, 2007, Cameron, 2006).

Cook and Rasmussen (2005) use the SWE (v1) in a conceptual paper on patient safety. They note that there is little if any research 'characterizing the location or movement of the OP of the system or reducing the size of the OP motions.' They also suggest that

further research is needed into the ‘factors influencing the marginal boundary location’. This research provides empirical evidence to address that gap in the current state of knowledge.

The extant SWE model has been developed by the synthesis of the literature to develop and add to the constructs and define their dimensions (Figure 9.1). The dimensions explored help to develop an understanding of the dynamic behaviour of hospitals as systems and the influence on patient safety. The specific contributions to knowledge relating to the SWE constructs are summarised in Table 9.2:

	Boundary
a.	The empirical evidence shows that the characteristics of the boundary construct display differences that help to explain the location and movement of marginal boundary and the prioritisation of competing requirements by decision makers.
b.	The inclusion of marginal zone boundaries for all failure boundaries increases the explanatory power of the model in relation to the resilience literature. The location and movement provides insight into the ‘buffer capacity’ available to be deployed to adapt to situations of perturbation or continuous stress. For example, where there is little or no buffer capacity it helps to explain why rapid actions are instigated to avoid a breach of a boundary.
	Operating Point
c.	Insights into the empirical data are derived through the synthesis of concepts from accident theory and systems thinking to explain how drift, tip, collapse and transition into failure can result from the dynamics interactions of the system characteristics.
d.	The empirical data presented on the movement and location of the OP provides evidence of the decompensation of buffer capacity. This builds on the extant literature, which only proposes such a phenomenon (Miller and Xiao, 2007, Cook and Rasmussen, 2005)
	Gradients
e.	The explanatory power of the gradient construct is extended by the synthesis of

	the literature to take account of the dialectic relationship between the ‘social structure’ and ‘agent’ derived from social theory. This allows insights into the dynamic interactive nature of the pressures exerted by the internal and external requirements and constraints.
f.	The gradients are not homogenous as implied in the literature (Rasmussen, 1997, Cook and Rasmussen, 2005, Miller and Xiao, 2007). Particular issues, such as control of infection or nurse working hours exert more pressure than other issues related to the same boundary gradient.
g.	The difference in pressure exerted by the gradients is associated with the timescale of the issue. The link between pressure and the timescale of issues associated with gradients is not considered in the extant literature.
	Structure and Feedback Loops
h.	SD provides a means to model and derive insights about the impact of the flow of work, the occupancy of the stocks, any changes in the designed flow and decision making.
i.	SD provides a means to explain the coupling between the parts and the types of feedback loops that influence the behaviour of the system.
j.	The use of SD expands the explanatory power of the model to show the vulnerability of a hospital when there is insufficient capacity to accommodate the net difference in flow rates

Table 9.2: Summary of the contribution to knowledge from the development of the SWE model

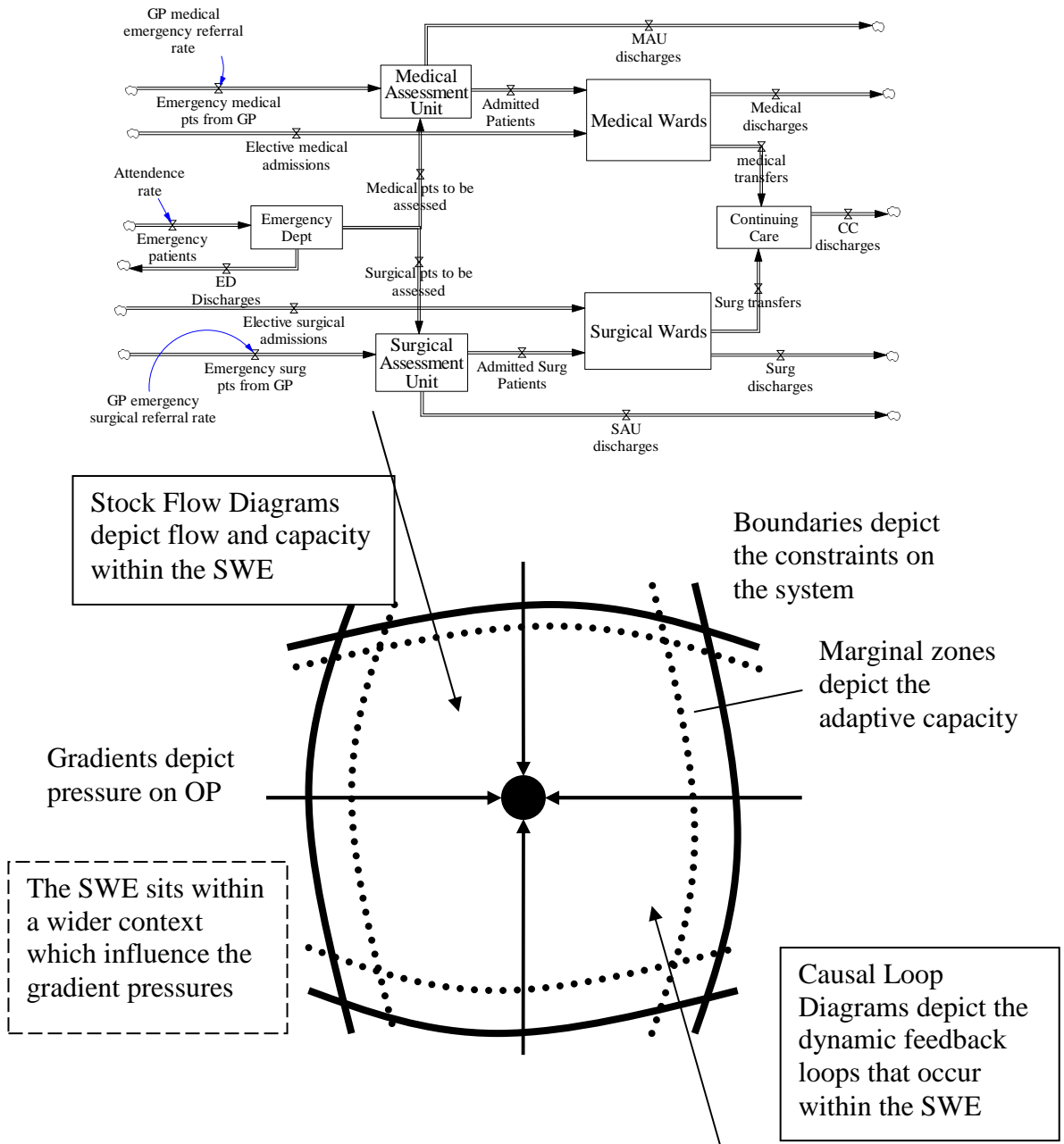


Figure 9.1: The constructs of the SWE v3 model

The SWE v3 model takes account of the presence of adaptive capacity in the marginal zones of the envelope. The presence or absence of that resilient capacity is considered next.

9.2.3 Resilience – the buffer of adaptive capacity

The concept of resilience explains how adaptive capacity is deployed to hold or return the OP within the SWE during periods of perturbation or continuous stress. The buffer capacity (marginal zone) depicts the ability of the system to adapt (take compensating actions) to keep the OP from breaching the failure boundary. As noted, the presence of ‘redundant’ resources added into the system to prevent failures can increase the complexity (Perrow, 1984, Sagan, 2004).

However, it is very clear from this research that the definition, measurement and monitoring of the boundaries of patient safety failure and unacceptable workload in the hospitals studied is weak in comparison to the other boundaries. Therefore, the resilient capacity (ability to adapt) relating to patient safety and workload is poorly understood. This is because the marginal zone and failure boundaries are not well defined and there are no measures of what constitutes adaptive capacity. Equally there is evidence that the target and finance failure boundaries were clearly defined, measured and performance managed. As there was limited adaptive capacity associated with some of the waiting time targets, the vulnerability of that boundary influenced decision makers.

Miller and Xaio (2007) argue that it is possible to measure the resilience of an organisation by identifying the location of the marginal zone boundary and the compensating mechanisms (buffer capacity) within the zone. For example, it is possible to measure the number of staff available to cover absences or the number of empty and staffed beds. As identified, not all the boundaries are homogenous or well defined. Building on the literature and the insights from this research, it is possible to propose a number of resilience indicators for each construct of the model (Table 9.3). Further work is needed to verify the indicators and to see if others are required. As Miller and Xaio (2007) suggest, being able to measure such indicators can provide a means to assess the resilience of a hospital.

Construct	Measurable indicators
Patient safety marginal zone	Number of patients with infections Number of beds available to isolate infectious patients Number of outliers Mortality and harm rates (indicators of harm and data collection methods to be established)
Unacceptable workload marginal zone	Number of unfilled gaps in front line staff rotas (Drs, nurses etc) Number of staff available at short notice to work Patient to Dr and patient to nurse ratio Acuity of patients (measure of how seriously ill patients are) Total walking distance of medical ward round Number of patients outlied from their expected ward
Target marginal zone	Performance against production / waiting time targets (ahead of target provides additional buffer leading to system resilience)
Financial marginal zone	Performance against budget (overspend creates less buffer leading to system vulnerability)
Stocks	Occupancy rate by hour, day, week, month and year Frequency of internal transfers for non-medical reasons (creating outliers)
Flows	Variation by hour, day, week and month Net difference by hour of flows in and out
Feedback	Planned design of the flow of patients compared to the actual Changes in the expected length of stay

Table 9.3: Measurable indicators of the constructs to provide an insight into state of resilience of a hospital

It is further argued that the ability to measure the proposed indicators can assist in overcoming the vulnerability of the current poor ‘visibility’ of the marginal boundaries and system dynamics of hospitals.

Miller and Xaio (2007) in their empirical study do not observe any ‘decompensation events’. Decompensation occurs when the buffer capacity is exhausted (Woods and Cook, 2006). This study has observed and presented data on both chronic and acute decompensation. The outbreak of the sickness virus and frequent, although short lasting shortage of staffed beds, are examples of acute decompensation. The events within CS 3 and the slowly rising need over three years to have more medical outliers in CS 1, is evidence of chronic decompensation. This research therefore provides evidence of decompensation events which adds to the literature.

The development of the Rasmussen (1997) model, and the subsequent application to undertake research exploring patient safety from a systems perspective, provides interesting insights into the characteristics and behaviours of the hospitals studied.

9.3 Characteristics of the case study hospitals

In this Section consideration is given to how the insights derived from using the SWE v3 model might be structured in a way to describe the system behaviours observed and their relationship to the risk of patient safety failure. The data gathered from the case studies provides some evidence about the performance of the hospitals in relation to the failure boundaries. This section draws on systems dynamics to help structure the discussion.

9.3.1 Insights into system archetypes

To explore how the system influences patient safety it is helpful to develop a theoretical framework that takes account of the points drawn from the literature. As set out above, it is proposed that the SWE (v3) model can provide insights into system characteristics, the resultant behaviours and the implications for patient safety. Meadows (2008) argues that taking a systems thinking approach seeks to understand the relationship between an ‘event’, the resulting ‘behaviour’ (e.g. oscillation) and the ‘structure’ of the system. She states that system structures that create patterns of behaviour are known as ‘archetypes’.

In developing and applying the SWE (v3) model three construct sets have been identified that interact to create the operational performance of the system (Figure 9.2):

- the constraints within which the system operates, depicted by the boundary construct of the SWE (v3) model (presented in Chapter 6);
- the pressures on the OP, depicted by the gradients (presented in Chapter 7);
- the dynamic interaction of demand, capacity and the decision makers within the hospital, depicted by the SD ‘structure’ part of the model (presented in Chapter 8).

The operational performance is depicted by the location and movement of the OP in relation to the failure boundaries. The ability of the system to avoid or recover from breaching a failure boundary is depicted by the marginal zones (adaptive capacity) of the SWE (v3).

9.3 System archetypes that may influence patient safety

9.3.1 One – Safe Practice

The outcome of safe practice is achieved where the interactions of the three construct sets create a dynamic equilibrium with the adaptive capacity. It is dynamic in that compensating actions are constantly required to maintain the safe practice. Without the actions an imbalance could occur, leading to an increased risk of interactions creating patient safety risks. In SD terms the balancing feedback loop/s dominate the system keeping the OP in a stable position.

9.3.2 Two – Drift towards Failure

There was clear evidence from the case studies of staff accepting small incremental changes to practice. For example, the gradually increasing numbers of medical outliers in CS 1, the movement of patients without their medical notes, the preparation of patients for theatre in offices, using day case beds for inpatients and the number of patient falls (Williams et al., 2009). This system archetype is that the performance standards are influenced by past performance (Meadows, 2008). Where past performance is perceived to be adequate, safe and required through practical necessity, then a reinforcing loop is generated that allows migration and violations to become the norm (Amalberti et al., 2006).

The case study data provides strong evidence of staff normalising over time to the continual small changes in processes and practice. The extant literature describes constant acceptance of small changes as a process of ‘drift’ (Dekker, 2011, Snook, 2000, Woods et al., 2010, Vaughan, 1996, Pidgeon, 1997). This research links the concept of drift to the interactions of the system constructs. Resilience theory helps to explain how the adaptive capacity can absorb the small incremental steps of practice change without any apparent safety concerns (Miller and Xiao, 2007). However, the chronic erosion of the adaptive capacity can result in the OP breaching the safety failure boundary.

There is irrefutable evidence that the acceptance of the incessant small changes results in the acceptance of higher risk operational practices. In Chapter 7 data is presented

which may be construed as the marginal patient safety boundary being shifted outwards when medical patients were outlied onto the gynaecology ward. The decision to outlie patients to that particular ward was only a small change from what had occurred previously, without any apparent safety risk. Yet the Lead Nurse for Control of Infection had indicated the potentially serious implications if infection from medical patients spread through the staff to new born babies and their mothers. The decision makers who chose to use the gynaecology ward for medical patients did not wish to increase the risk of infection to staff and babies. However, the cumulative effect of the competing and interacting system construct sets left the managers little alternative. When justifying the decision, managers regarded the risk of not being able to admit emergency patients (pressure from the target gradient) as greater than the small change in the policy of outlying patients to the gynaecology ward (lack of pressure from the patient safety gradient). The practical requirements of the situation meant that a further small step in the process of drift to failure occurred.

Analysis of the situation, using the SWE (v3) model facilitates the identification of the differences in visibility, buffer capacity and pressure from the target failure boundary, in comparison to the patient safety failure boundary, influenced decision makers. The stark examples of operational practice slowly drifting to failure can be conceptualised as the system being subject to periods of continuous stress, which pushes the OP towards the patient safety failure boundary. In the case studies, periods of continuous stress were observed. For example, the evidence of the constant variability of demand of patients requiring admission was driven by a number of factors. These include changes in the morbidity of the population (e.g. rise in breathing problems in cold weather), the waiting time targets and the availability of staff and resources to treat them. The response of decision makers to the pressures, constraints, workload and capacity dynamics influenced the degree of continuous stress experienced by the system. As noted in Chapter 8, the inability to accommodate the accumulation of the net difference in rates of flow, created the conditions where the systems drifted towards failure as the number of medical outliers increased.

CS 3 provides substantial evidence of a long term drift to failure. The second Public Inquiry into the wider influences on the hospital will not be published until 2012. However, the first inquiry report includes many accounts of how the need to save money and meet targets influenced the decisions to reduce staffing. The accounts from

clinical staff show a process of drift and normalisation. The gradual decline in standards of treatment of patients is a clear example of the drift to failure. In CS 3, it is argued that conceptually the OP clearly breached the patient safety failure boundary on repeated occasions, often without staff noticing.

9.3.3 Three – Tip towards Failure

Cook and Rasmussen (2005) argue that the SD of an organisation can rapidly change when the system reaching a ‘tipping point’. They give the example of a steam boiler where the dynamics associated with steam are very different to water although the change in temperature between the two states is small. Sudden changes in the dynamics were found in the CS 1 and 2. The results present instances of departments and hospitals reaching a ‘tipping point’. At the tipping point the system changes from being reasonably efficient to where major inefficiencies were experienced.

The system behaviour is where the system oscillates around the load carrying capacity of the system. Balancing feedback loops are required to act quickly to counter the reinforcing feedback loops and avoid a system overshooting the capacity to meet the demand (Sterman, 2001). When time delays occur in the feedback loops the system can be working above their capacity and reach a ‘tipping point’ where performance deteriorates rapidly. When the balancing feedback loops take effect, the system returns to a stable state, where further demand is experienced and the pattern is repeated.

The examples found in the case studies relate to the capacity to deal with fluctuations in demand for services. The incidents arose as a result of either continuous stress (just one more patient) or a perturbation (sudden rise in demand, the loss of bed or staff capacity) to the system. Key bottleneck departments, such as the MAU / EMU, were evident in the flow of medical emergency patients. The bottleneck departments were often unable to accommodate newly arriving patients. Both CS 1 and 2 provide clear evidence of reaching a tipping point where the hospitals moved rapidly from having loosely coupled to tightly coupled departments and hospitals. One additional patient arriving in the MAU, when it was full, created a diversion of GP admissions to the ED or the outlying of medical patients into surgical beds. The diversions and outlying are examples of instigating balancing loops to return the bottleneck department to a position of being able to meet the demand. However, as clearly demonstrated in Chapter 8, a consequence of instigating the balancing loops for the MAU is that the hospital tipped into a tightly

coupled state. The number of reinforcing feedback loops in the hospital system then increased with detrimental effect on patients and staff working conditions.

The tipping point for the system is often visible, for example when the destination of patient flow changes. The SWE (v3) provides a way to explain the longer term interactions that allow such a tipping point to be reached. The constraint and pressure over the longer term from the financial boundary seek to maximise efficiency and therefore, not to invest in additional capacity. The financial pressure combined with the short term constraints and pressure from the various waiting time targets and the variation in demand. The result was the inability to accommodate the imbalance of flow rates in and out of the hospital. When bottleneck departments are unable to accommodate the accumulation of patients created by the disequilibrium in flow rates, then the system is at risk of tipping into a reinforcing cycle of inefficiency. Data from interviews and observation provided repeated examples of doctors chasing after patients, breakdown in communication between clinical teams, delays in investigations and treatments and elderly patients being moved in the early hours of the morning.

As noted, accident and systems theory argue that the number of interactions increase as organisations become tightly coupled (Perrow, 1984, Dekker, 2011). The complexity of the situation therefore escalates. In CS 2 the ED was observed when it experienced a tipping point. The department rapidly changed from a position where staff were able to manage the flow of patients to the point where patients were being held in the corridor, the waiting room and were inappropriately using the resuscitation room. The ED changed from a smooth running unit to being on the edge of chaos. It should be noted that sudden change in the dynamics of the system can be short lived or last several days (Wears et al., 2008).

There was clear evidence that staff had normalised to the increased frequency of oscillating around the capacity limit. The resultant inefficiencies were regarded as inevitable by many interviewees. The implications for patient safety did not appear to take a high priority. Analysis using the SWE (v3) model would suggest that the lack of visibility of the patient safety failure boundary made it possible for staff to normalise to the tipping point. The constraints and pressures on the system, and lack of capacity to deal with perturbations or continuous stress, become accepted as the norm. The SWE (v3) model approach argues that short term capacity problems can be conceptualised as

the exhaustion of the buffer capacity (Woods and Cook, 2006). Where that exhaustion does not last long it is termed 'acute decompensation; defined as the short-term exhaustion of compensatory mechanisms' (Miller and Xiao, 2007)

9.3.4 Four – Collapse towards Failure

The sudden and unexpected breaching of the patient safety failure boundary by the OP was evident in the outbreak of the Norovirus in CS 1. Such a situation is an example of what can be described as 'immediate decompensation' of the adaptive capacity, or a collapse towards failure. The system behaviour is where reinforcing loops drive exponential growth until the sudden collapse of the performance of the system (Sterman, 2000). This can be caused by an external disruption or the gradual erosion of adaptive capacity, leading to a position where the system is quickly overwhelmed and suddenly collapses. Unlike the oscillating archetype, it takes time to rebuild the adaptive capacity; hence the term 'collapse' is used.

The Norovirus outbreak overwhelmed the normal infection control practices and precautions with the rapid spread of the virus across a number of wards. As noted, rapid action to contain and rectify the situation took place, although it took many days to recover. The perturbation also impacted on the operational management of the whole hospital. The financial gradient was relaxed, to allow additional staff to be employed. However, there was an immediate capacity problem due to the beds being closed to control the spread of infection. The decision not to cancel elective admissions in this situation is a stark example of the target gradient taking precedent over other pressures and constraints. This exacerbated the pressure on the OP moving towards the patient safety and unacceptable workload failure boundaries. Keeping the production pressure on the OP meant that the reinforcing loops continued to dominate the system.

The sudden increased numbers of admission in CS 1 after this Christmas period is an example of the system collapsing towards failure. The hospital capacity was rapidly overwhelmed and experienced immediate tight coupling as the system went 'solid' (Cook and Rasmussen, 2005). There was also a collapse in the hierarchy with Directors becoming involved as porters and in the management of individual patient movements. The rapid rise in the number of medical outliers was undertaken with little, if any regard, to the safety implications. Whilst some elective cases were cancelled, bank staff deployed and extra money spent, the production pressure dominated. This is a

clear example of the OP moving rapidly through the marginal zone at the intersection of the patient safety and unacceptable workload boundaries. The adaptive capacity was immediately exhausted as the medical outliers rose to over one hundred. Although there did not appear to be any reported major failures in patient safety, the literature suggests that error and harm may have occurred but was not noticed or reported (Tucker and Edmondson, 2003, Vincent, 2007, Waring, 2005, Godlee, 2007, Leape, 2002, Health Care Commission, 2009, Tucker, 2004, Healthcare Commission, 2009a, Healthcare Commission, 2009b).

9.3.5 Five – Transition towards Failure

The fifth system behaviour is proposed as means to describe the combination of interactions that lead to the system performance becoming unsafe. A characteristic of complex systems is the nonlinear interrelationships which can lead to unexpected consequences (Dekker, 2011). This system behaviour is the dynamic interaction between the reinforcing and balancing feedback loops that create a variety of potential system behaviours.

Some of examples found in the case studies arise from a combination of underlying drift with chronic or acute decompensation (Miller and Xiao, 2007). However, there is evidence that there are numerous interactions that lead to safety failure. There is the case from CS 2 of the infected patient with an underlying heart condition being held in the ED and not being treated in a timely manner. The combination of strong pressure to contain the spread of infection, the immediate shortage of single rooms, the rota arrangements for junior doctors and the competency of staff and timing of review by a senior doctor, all contributed to the transition to failure. This provides an excellent example of compensating actions (isolate infected patients; reduce working hours of doctors) interacting with other factors in potentially unforeseen ways to transition towards failure.

The concepts of latent conditions (Reason, 1997) or the incubation of disaster theory (Turner and Pidgeon, 1997) provide some explanation for the transition to failure. However, these explanations are limited as they focus on understanding the component conditions and how defences against component failure do not prevent an accident. The proposition of ‘transition towards failure’ adds to the explanatory power of extant literature by describing the interactions between the construct sets. Reason’s (1997)

argument that latent conditions combine with local operating conditions is not sufficient as it advocates a linear cause effect. The system behaviour of 'transition to failure' takes the dynamic feedback between the 'structure' and 'agent', found in social theory (Held and Thompson, 1989, King, 2004, Lane, 1999, Lane and Husemann, 2008a, Thompson, 1989, Giddens, 1984), into account.

The system archetypes presented have been considered at a hospital level. It is however feasible that such system structures and resultant behaviours occur in the working of the micro clinical teams, departments and the wider hospital (Mohr et al., 2004). This is evidenced in the example from CS 2 where the infected patient was held in the ED. There were a series of interactions at the different levels within the hospital. Such a view fits with the nature of complex systems (Cilliers, 1998) and the idea of nested cycles of adaptive capacity (Holling, 2001). However, detailed exploration of this idea is beyond the scope of this research.

9.4 Proposals for improvement

The pragmatic critical realist position of this research leads to a number of possible proposals about how to reduce the risk to patient safety. The discussion emphasises that a major influence on the behaviour of the hospital are the dynamics generated by the interaction of the construct sets of constraints, pressures, demand, capacity and the decision makers. The risk of systems experiencing any of the four archetypes of drift, tip, collapse and transition towards failure may be reduced if certain actions are taken.

There are opportunities to change the system structure of interacting feedback loops to avoid the problematic system behaviours (Meadows, 2008). The first point of intervention in healthcare is to prevent the drift associated with the continuous production pressures. The second is to avoid the shift in complexity associated with moving from a loose to tightly coupled system.

The first proposal addresses the system behaviour of 'drift'. The chronic erosion of adaptive capacity often occurs as a result of the production and financial pressures on the hospital system. It is very clear from both the extant literature and this research that the lack of definition and consequent visibility of the patient safety and unacceptable workload boundaries are key in allowing 'drift' to take place. Safety performance

standards have to be made clear and externally benchmarked to the best in class, rather than compared only to previous internal performance.

The resources allocated within the case study hospitals to define, measure and manage the financial and waiting time target performance was considerable. It is harder to define and measure the other boundaries. However, as suggested in Table 9.3, there are indicators that could be measured. The theoretical framework of the SWE (v3) does provide a means of visualising the balance between the competing dynamics that has to be achieved. However, this requires better definition of the boundaries to be made, understood and performance managed. The proposition is therefore made that the more hospitals clearly define, measure and performance manage patient safety and workload standards against external benchmarks, the risk to patient safety will reduce.

The second proposal addresses the ability to manage the disequilibrium in the rates of flow into and out the hospital is a key point of intervention. The vulnerability of the system to creating potentially unsafe feedback loops is linked to the bottleneck in the flow of patients that occurs in MAU / EMU. When there is insufficient capacity to accommodate the accumulation in the net difference in rates of flow, the coupling between parts of the hospital increases. The overflow of patients into the surgical department occurs with many consequences for the hospital, staff and patients. It is important to note that the isolation stock (capacity to accommodate infectious patients) has also to be able to manage the net difference in flow rates.

It is easy to propose additional capacity but most hospitals face financial constraints and pressure which prevent additional expenditure. However, failure to avoid moving to a position of tight coupling is a false economy. Not being able to accommodate the demand flow can create major inefficiencies and risks to patient safety, consequently increasing the costs of treatment (Blackstone, 2004, Colwyn Jones and Dugdale, 1998, Goldratt and Cox, 2004). Therefore, emphasis has to be placed on the effectiveness and efficiency of the processes of assessment, investigation, treatment and transfer of patient from the MAU facility.

There are a number of operational management techniques that can assist in managing the flow of patients. These include putting the senior expertise as close to the beginning of the process of assessment, investigation and treatment as possible (Seddon, 2008).

This can reduce the time delays that occur when junior doctors have to handover information to senior colleagues for decisions to be made. As evidenced by the case studies the handover and decision making is often done during ward rounds, with the resulting batching of work. As was evident from CS 1, batching and the delay in processing add to the inefficiencies and slow down the rate of flow out of the hospital.

A technique derived from the theory of constraints is to set up a 'buffer management' process to protect the physical bottleneck (Gardiner and Blackstone, 1998, Goldratt and Cox, 2004). Amongst other benefits, buffer management can provide an early warning to decision makers of times when the MAU may not be able to accommodate the net difference in flow rates. Such a warning can allow compensating actions to be taken prior to an overflow of patients. The key point is that the avoidance of medical outliers and the diversion of patients from MAU to ED is a critical leverage point in improving the system behaviour and reducing the risk of unsafe practice. Therefore, the proposition is made that the more a hospital can absorb the net difference in flow rates, the risk to patient safety created by tight coupling feedback loops will be reduced.

The implications of these proposals for policy makers and practitioners are explored further in Chapter 10.

9.5 Summary of the contribution to knowledge

The contribution to knowledge is in three areas. The first is the development of the literature on resilience and the SWE model in particular. Second, how the system characteristics influence patient safety and third, the proposals for improvement.

The first area of contribution is that the research revealed the requirement to develop the Rasmussen (1997) SWE model. The model has been extended by the addition of a 'target failure' boundary and the application of SD to gain insights into the dynamics that occur when demand and capacity interact with decision makers. The resulting SWE v3 model provides a new way to gain insights into the system behaviours of the hospitals studied. The contribution to knowledge is in both the development and the application of the model to consider patient safety from a resilience and dynamic rather than linear perspective.

Second, insights are gained and articulated that develop the understanding about the interactive and emergent nature of some of the system behaviours in NHS hospitals and how they influence patient safety. Those insights are viewed from the perspective of regarding patient safety as an emergent property of the dynamic interactions represented by the constructs used in the SWE (v3) model. Using the model provides a means of exploring the interactions and decision making that takes place. The analysis of the case studies using the SWE (v3) model develops the theoretical framework of the system construct sets and archetypes presented in Figure 9.1 and Section 9.3. The framework adds a different perspective to the extant patient safety literature. In particular the analysis using the synthesis of concepts from systems thinking, resilience and accident theory provides a new approach to considering how the 'system' influences patient safety.

The third area of contribution is the propositions for improvement set out in Section 9.4. The proposals are derived from the analysis of the case studies using the SWE (v3) model. A working assumption, derived from the literature, is that overcrowded hospitals are associated with patient safety failure. Therefore, the key outcome of this exploratory research is that tight coupling and the lack of buffer capacity within hospital systems, makes them vulnerable to breaching the patient safety boundary. Hence the proposition is made that the more a hospital can absorb the net difference in flow rates, the risk to patient safety created by tight coupling feedback loops will be reduced.

The measurable indicators of resilience presented (Table 9.3) develops the systems resilient approach to patient safety. This is a development of the literature (Miller and Xiao, 2007), which suggests that it is possible to measure the resilience of an organisation by identifying the location of the marginal zone boundary and the compensating mechanisms (buffer capacity) within the zone. The indicators proposed for each boundary allow an assessment to be made about the resilience of the hospital and the balance achieved by decision makers between the need to meet production pressures and to protect patients and staff. The proposition is therefore made that the more hospitals clearly define, measure and performance manage patient safety and workload standards against external benchmarks, the risk to patient safety will reduce.

9.6 Summary

This chapter discusses the findings from the case studies and makes proposals. The research objective has been explored using a system resilience model (SWE v3) developed from the literature and empirical research. Three system construct sets of the hospitals examined have been identified. The interactive construct sets create five possible system archetypes, four of characterise conditions where patient safety decreases. Proposals for improvement are made which focus on helping hospitals to not drift or tip towards failure. The contributions to knowledge are summarised as being first, to the development of a new theoretical framework to better explain how the system level influences patient safety. Second, the proposals for improvement in the operational management that impact on patient safety, and third the development of the SWE model.

The implications, limitations and potential opportunities for future research are discussed in the next chapter.

Chapter 10 - Conclusions, Implications and Limitations

10.1 Introduction

This chapter summarises the conclusions about using the SWE (v3) to gain insights into the systemic characteristics of NHS hospitals in England and how they influence patient safety. The implications of these conclusions for theory, policy and practice are presented. The limitations of the research are noted and proposals for future research are made.

10.2 Summary of Conclusions

It is argued in Chapter 2 that the extant patient safety literature highlights the need to consider the ‘system’ contribution to patient safety. However, there is little application of systems thinking in developing the explanation of how that influence occurs. This research uses a conceptual model that addresses safety in complex socio-technical systems (Rasmussen, 1997). The results provide evidence about the construct sets of NHS hospitals and how they interact to create system behaviours that influence the safety of patients. There are conclusions drawn in Section 10.2.1 about the adequacy and the development of the Rasmussen (1997) model. In Section 10.2.2 conclusions are presented from the research about patient safety in the case study hospitals.

10.2.1 Conclusions about safe working envelope model

The SWE (v1) (Rasmussen, 1997) provides a model that takes account of the constraints and competing pressures that influence complex socio-technical systems. The model depicts the conceptual idea of a resilient system as being one that can maintain the OP within the failure boundaries of the envelope. However, in Chapter 5 it is argued that for the model to be applied to the context of NHS hospitals in England, it needs development.

The first development is the additional boundary of ‘target failure’. The main evidence for the need of an additional boundary comes from the examination of policy documents from the Department of Health. The documents place considerable emphasis on achievement of waiting time targets (see Section 4.2.1), which is supported by the extant literature. It is concluded that the addition of this boundary extends the ability of

the model to capture the complex influences created by patient demand. The influence of patient demand is otherwise difficult to take into account when using SWE (v1) (Miller and Xiao, 2007).

The SWE (v1) model is used in the resilience and safety in healthcare literature (Miller and Xiao, 2007, Amalberti et al., 2006, Cook and Rasmussen, 2005). However, it is argued that the original model is weak in taking account of production pressures relating to targets, 'structure' and 'feedback' (Sterman, 2000). To overcome these weaknesses, in the second development, aspects of SD are used to extend the model (see Section 2.5.1). The ability of the SWE (v3) model to take account of production pressures, structure and feedback is evidenced in Section 5.3 and Chapter 8. The contribution to knowledge resulting from the use of SD as part of the SWE (v3) model is presented in Table 9.2.

The third development is using the concept of 'cross scale' interactions (Woods, 2006) being applied to the OP (see Section 3.3.1). The downward scale interactions assumed in the original model depict the gradients moving the OP away from the failure boundary (Rasmussen, 1997). By using upward scale actions it is possible to assess actions by staff that can create pressure to move the OP towards a boundary (see Figures 7.2, 7.3 and 7.8). The use of 'cross scale' interactions develops the model from depicting a linear pressure on the OP, to where it is possible to take account of the dialectic interactions between the gradient downward pressure and the response of staff. It is concluded that this development boosts the ability of the model to take account of the dynamic interactions within the system being studied.

The fourth development is the explicit use of marginal zone boundaries associated with each of the four envelope boundaries. The findings suggest that the characteristics of the failure boundaries show some differences from each other (see Tables 6.4 – 6.7). These include differences in the visibility, movement and location of the marginal zone boundary and size of the buffer capacity. It is concluded that the differences provide an insight into the pressures from the gradients associated with each boundary. This insight helps to provide an explanation beyond that found in the extant literature as to why decision makers pay more attention to the financial and target failure boundaries than the others.

It is concluded that many of the concepts derived from the accident theory literature can be incorporated within the SWE (v3) model (see Section 3.4). The development of the model allows for the systemic concepts of coupling and feedback loops to be included (see Table 8.2). It is concluded that the development of Rasmussen's (1997) SWE (v1) model does increase the explanatory power of why the pattern of systemic behaviours arise and influence patient safety in the case study hospitals. The implications of the model development for wider application are considered below.

10.2.2 Conclusions about the systemic characteristics influence on patient safety

The use of the SWE (v3) provides insights into the systemic characteristics of the case study hospitals. Any model seeks to depict reality but is limited. Therefore, the insights are not comprehensive. However, it is argued that the model does allow the identification of important construct sets and helps to explain the patterns of behaviour of hospitals, which may influence patient safety. Woods et al (2010) argue that safety is an emergent property of complex systems. Safety emerges from the adaptive capacity of the many actors in the system. The SWE (v3) provides a means to explain something of the complexity of the dynamic interactions that are associated with patient safety in the case study hospitals. It is concluded that some systemic characteristics of the hospitals arise from the interaction of the construct sets analysed.

The argument of this thesis is that the dynamic interactions related to the three construct sets often create unpredicted feedback loops. Five system behaviour archetypes are identified, four of which have a negative influence on the safety of patients. Reinforcing feedback loops create pressure on the OP. A key finding of this research is that the competing pressures on the OP create the situation where the dynamic interactive complexity of systems can escalate either slowly (drift towards failure) or rapidly (tip, collapse or transition towards danger). This is demonstrated by the number and type of feedback loops that are created as the hospitals become tightly coupled.

When hospitals have a high level of occupancy they appear to be more vulnerable to perturbations and continuous stress which move the hospital into being tightly coupled. It is argued that when the hospital becomes tightly coupled, then the resilient adaptive capacity of the staff to keep patients safe is stretched, which increases the risk of harm occurring (Woods et al., 2010). The vulnerability due to tight coupling provides a

greater insight than the suggestion of an 85% occupancy to avoid overcrowding (Bagust et al., 1999).

When harm occurs to patients, it is normal in the NHS for an investigation to seek out a root cause through a process of linear analysis and hindsight bias (Dekker, 2006). As the likely cause of harm is the combination of dynamic interactions, such an approach does not provide the full explanation (Dekker, 2011, Hollnagel, 2004, Dekker, 2006). Therefore, a conclusion of this research is that a systemic understanding of the hospital is required to address the fundamental vulnerabilities to patient safety failure found in the case study hospitals. The argument is that the vulnerability and resilience of the system is related to the adaptive ability to maintain the OP within the boundaries of the envelope at times of problematic system behaviour. The SWE (v3) provides a mechanism to conceptualise insights into the adaptive or buffer capacity associated with the envelope boundaries and stocks in the system.

When the marginal zone boundary is either located or moves close to a failure boundary, then the size of the buffer capacity that can be deployed to adapt to perturbations or continuous stress is reduced. For example, the evidence from the findings suggests that the location of the marginal target zone boundary is close to the failure boundary. The conclusion is drawn that the corresponding lack of buffer capacity, high visibility and performance management of the target failure boundary helps to explain the pattern of high pressure from the associated gradient. The dimensions of the target boundary also assist in understanding why consequent compensating actions initiated by staff often appear to take priority over other competing requirements. For example, admitting patients within four hours took a higher priority than finding the appropriate location for treatment.

The extant literature suggests that the visibility (more easily defined, measured and monitored) of the production related boundaries (financial and target) is one of the main reasons why they are given greater priority over the workload or safety failure boundaries. This research supports this conclusion. However, it is argued that the location and movement of the marginal zone boundary depicting the buffer capacity, adds to the explanation (see Section 6.3.4).

Building on the work of Miller and Xiao (2007), there are some measurable indicators of the buffer capacity that can be proposed for each boundary (see Table 9.3). Further work is required to substantiate the proposals. However, the use of the marginal zone boundary does add to the explanation about the resilience by being able to propose what constitutes buffer, or resilient adaptive, capacity for each boundary.

A key conclusion is the need to have sufficient buffer capacity in the medical department stocks, especially the MAU, to prevent the flow of patients diverting into surgical department stocks. The need to borrow buffer capacity is shown to create conditions of tight coupling and reinforcing loops that increase the complexity of the system for staff. It is concluded that when reinforcing loops dominate the dynamics, the consequent system behaviours moves the OP towards the patient safety failure boundary.

10.3 Implications

There are implications of the outcome of this research for theory, policy and practice which are presented in this section.

10.3.1 Implications for theory

This research takes an exploratory and theory building approach (Eisenhardt, 1989, Meredith, 1993). It is argued that the extant patient safety literature does not have an adequate theoretical model with sufficient power to explain how the dynamic interactions within healthcare systems influence patient safety. The gap in knowledge is addressed by using concepts from resilience engineering and SD literature to extend and apply the SWE (v1) model.

The contribution to patient safety theory is to suggest that the construct sets used in the SWE v3 model of hospital systems combine, interact and create five system behaviours that influence safe performance. The extant literature, explored in Chapter 2, lists numerous organisational factors that contribute to patient safety, but is limited in exploring the implications of the systemic interactions between the many factors. The Swiss Cheese theoretical model appears to be most popular in the literature to explain why patient safety failures occur. It is argued that such a model takes a linear and

barrier approach to safety and is therefore inadequate when seeking to explain systemic influences on patient safety.

The workload of staff at an individual level is explored in the literature from a patient safety perspective (Reason, 1990). However, the SWE (v3) model provides a broader perspective. The model allows staff workload to be viewed within a systemic perspective with competing goals, tight coupling and dynamic feedback. The drive for efficiency in systems is accounted for in the extant literature (Reason, 1997). This empirical study depicts the implications of the requirement to reduce costs and increase productivity in terms of the location of the OP and the impact on the buffer capacity of the stocks and boundaries.

There is an implication for theory in the link between staffing workload and patient safety. There are system level implications on the operational and staff capability of hospitals arising from high levels of patient demand. For example, the patient safety and staff workload implications of patients being outlied from medical wards have not been well researched. It is suggested that patient safety theory builds more explicit links with operations management and systems theory to address such issues.

As noted, the ability of the SWE (v3) to incorporate concepts from accident theory, adds to the explanatory power of the model in considering the safety issues that arise in socio-technical systems. There are also implications for theory in considering patient safety from the perspective of the resilience of the hospital system. The SWE (v3) provides the conceptual model for considering the buffer capacity required to maintain the OP within the constraints of the envelope boundaries. This research has added to the knowledge about the system resilience of the hospitals investigated through the proposed indicators of buffer capacity.

An implication for theory arises from the combination of concepts from social theory, SD and resilience engineering. Patient safety theory is multifaceted with theories derived from many academic disciplines (Vincent, 2010). It is often the synthesis of concepts from different viewpoints that create opportunities to develop theory. It is argued that the contribution to knowledge from this research results largely from the synthesis of ideas from different disciplines to develop and apply the SWE (v3) model.

The final implication for theory is the suggestion that the SWE (v3) may have wider application than the case study hospitals. It is suggested that the model can be adapted and applied to any system that has the combination of financial requirements, production targets, staff undertaking work, standards of expected performance and a flow of work through the system.

10.3.2 Implications for policy

There are implications for policy makers from the outcomes of this research. NHS policy in recent years has placed greater emphasis on patient safety as part of a wider quality agenda (Department of Health, 2008b). However, largely absent from that policy is the recognition of staff workload on patient safety and the influence of competing dynamics created by conflicting policy goals. Targets are set for specific types of patient safety failures, such as infections and high risk procedures (Department of Health, 2009b, Department of Health, 2010d) which indicate a component rather than a systemic approach. There is little if any recognition in the documents that other policies associated with waiting times and financial management can interact to reduce adaptive capacity and increase the risks of harm to patients. The application of systems thinking, and an understanding of what constitutes system resilience, will assist policy and regulation in creating the conditions for greater adaptability by front line staff to avoid patient harm.

With the current situation of the UK economy, the financial savings required across the NHS means that there is a renewed emphasis on the 'faster, better, cheaper' approach. Research indicates that such an approach can create conditions where deviance is normalised (Vaughan, 1996). The emphasis on efficiency can result in the buffer capacity being regarded as 'waste' and removed (Radnor and Walley, 2008, Smart et al., 2003, Lawson, 2001, Woods et al., 2010). The removal of so-called waste is often thought to be of practical necessity to meet the efficiency requirements. When no safety failure is evident, the action is thought justified and often repeated. However, the implications are that future costs will be higher, when things do go wrong (Woods et al., 2010), as evidenced by CS 3.

The current most popular theory used by policy makers, regulators and practitioners to explain why accidents occur in healthcare is the Swiss Cheese model (Reason, 1997). As noted above, this is not sufficient to explain the systemic interactions that occur in

complex systems such as hospitals. The use of the SWE (v3) model allows other insights to guide policy, regulation and practice. A key implication of adopting a different theoretical model will be a change in policy from the 'root cause' linear investigation of patient safety incidents referred to above to a more system based method (Dekker, 2006).

In summary the implications for policy makers in the NHS are:

- To take a system based approach to policy relating to performance management and targets to understand the interactive implications.
- To place patient safety as a *consistent* core organising principle in key documents like the Operating Framework. Fundamental to achieving safety is the assessment of the resilient capacity of healthcare services to sustain safe practice in the face of high workloads and shrinking budgets.
- To assess the impact of policies and subsequent contracts for providers on staff workload and the consequent implications for patient safety.
- To revise the use of linear based models of investigating patient safety failures and adopt more system based methods (Dekker, 2006)

10.3.3 Implications for practice

The key implication of a system resilience approach to patient safety for hospitals is for safety to be embedded in the strategic and operational management of the institution. Hospitals are hazardous places for patients yet the vast majority are kept safe by the adaptive capacity and professionalism of the staff (Berwick, 1991, Woods et al., 2010). Failures are to be expected and learnt from as insights into the systemic characteristics of the hospital. Blaming individuals should be the exception rather than the expectation (Dekker, 2007).

The case study hospitals have invested considerable resources in being able to define, monitor and manage their financial, waiting times and infection control target performance. The implication of this research is that a comparable resource should be deployed to achieve the same level of visibility and attention to staff workload and patient safety. The SWE (v3) does provide a model from which to build indicators for the location of the OP, marginal boundaries and assess the size of the buffer capacity of

the stocks and marginal zones. Table 9.1 presents the proposed indicators that could be measured.

The operations management of the hospitals can be strengthened from using the ideas from SD about stocks, flows and feedback. A more sophisticated understanding of the occupancy and flows of work by hour of the day, day of the week and season is required to assist decision makers to balance and trade off the competing demands. These actions are required to avoid the system behaviours that lead to danger. Early warning of the depletion of buffer capacity in the stocks that are critical to the flow of emergency patients will assist in preventing the hospital reaching a situation of tight coupling with increasingly complex interactions. The stocks of the ED and MAU have to be sized and managed to accommodate the net difference in flow rates. The batching of medical work can be reduced to improve the flow of patients through the support departments of the hospital and therefore reduce the length of stay.

In summary the implications for practitioners at different levels in NHS hospitals are:

Trust Board:

- To define, measure and monitor key indicators of patient safety and buffer capacity. These indicators should include the impact on patients of high staff workload and the safety of outlying patients due to overcrowding of the emergency bed capacity.
- To benchmark safety performance against the best in class in order to avoid 'drift' in standards and normalisation by staff.
- To revise and simplify the policy on reporting to encourage greater openness about actual and potential safety problems.
- To revise the root cause approach to serious incident investigations to a more system based approach which seeks to understand why actions were taken and how the interactions occurred.
- To monitor the frequency of borrowing bed capacity from outlying patients or using day case facilities for inpatients; the length of stay by patient groups and the admission and discharge patterns.
- To size the bottleneck departments to be able to accommodate the net difference in flow rates.

- To increase knowledge about patient safety, human factors and accident theory and their relationship to the operational management of the hospital and contractual agreements.

Department

- To monitor and manage the flow rates to avoid having to outlie patients.
- Seek to improve the continuity of care by clinical teams and avoid non-medical movement of patients.
- Encourage reporting of system issues that create problems for clinical teams.
- Develop key indicators of patient safety and staff workload.
- Develop expertise in staff to take a system based approach to investigate incidents

Clinical Teams

- To increase knowledge about human factors and the causes of error and revise ways to investigate serious incidents.
- Develop an awareness of a safe working envelope for the team and the ability to assess performance in relation to the failure boundaries.
- Report concerns about safety and related system and capacity issues.
- Be involved in any decisions about non-medical patient movement.

10.3.4 Contribution to knowledge

In summary, and as noted in more detail in Section 9.5, there are three main contributions to the extant knowledge.

First, the analysis of the case studies develops the model of three interacting construct sets. The interactions of the constructs create emergent behaviours, some of which are problematic for patient safety. Five archetypes of system behaviour are identified, four of which increase the risk to the safety of patients. The SWE v3 model, construct sets and system behaviour archetypes adds a different perspective to the extant patient safety literature. In particular the analysis using the synthesis of concepts from systems thinking, resilience and accident theory provides a new approach to considering how the 'system' influences patient safety.

Second, propositions are made about how to reduce the risk associated with the system archetypes. A key outcome of this exploratory research is that tight coupling between the parts of hospitals and the lack of buffer capacity makes them vulnerable to patient safety failure. Hence the proposition is made that hospitals must be able to accommodate the net difference in rates of flow into and out of the hospital to avoid medical patients being accommodated on surgical wards. The case studies show evidence of the gradual deterioration of standards of care that can occur when hospitals are under continuous pressure to meet targets and save money.

The proposition is made that the safety standards have to be made more explicit and judged against high performing organisations and not internal past performance. Building on the extant literature, which suggests that aspects of resilience can be measured, a range of variables are proposed (set out in Table 9.1), to measure the adaptive capacity of hospitals relating to finance, targets, staff workload and patient safety. It is argued that the measures provide a means to assess the resilience of a hospital and the balance made by decision makers between production pressures and the need to protect patients and staff.

Third, is the contribution to the patient safety and resilience literature through the incremental development of the SWE model. The development of the model through the synthesis of the literature provides greater explanatory power in the exploration of some of the interactive components found in NHS hospitals that influence patient safety. The case study evidence adds to the literature that has used the SWE model by providing data about phenomenon not previously explored empirically. For example, both the chronic and sudden exhaustion of adaptive capacity to prevent patient safety failure is found in the case studies, provides a deeper understanding as to how that occurs.

10.4 Limitations of the research

Whilst it is argued that new knowledge is derived from the analysis of the rich data collected during this research, it is necessary to set out the limitations of the methods employed and conclusions drawn. This section addresses the limitations and how some of the limitations were handled during the research process.

The limitations of case study design are addressed in the literature (Stuart et al., 2002, Voss et al., 2002) although some are disputed (Flyvbjerg, 2006). A number of strategies and research methods, set out in Chapter 4, are used in the case study design, data collection and analysis to ensure the validity and reliability of the research outcomes (Yin, 2003, Eisenhardt, 1989). This has included the testing of conclusions with other academics, with managers and clinicians from the hospitals and the wider healthcare community. The data has been analysed, reduced and presented within the thesis and compared to the extant literature. The mixed methods of data collection provide a large volume of rich material, which can be overwhelming and lead to complex and unhelpful theories. The SWE (v3) model outlined in Chapter 5 is used to provide the framework for the data collection, analysis and presentation to focus the research.

Whilst any model has limitations this model is chosen to reflect the dynamic competing pressures that arise in complex socio-technical systems such as hospitals. However, it is not claimed that the model provides an exhaustive means to explore the characteristics of the hospitals. It does provide for a high level analysis of how the competing dynamics of finance, staff workload, waiting time targets and the flow of work can influence the safety of patients. As noted above, a limitation is the ability to know to extent to which patient safety is influenced by the interactions as many safety failures are either not noticed or not reported. The assumption is made that overcrowded hospitals increase the risk of patient safety failure. While this assumption is supported by the literature the limitation is that specific threats to safety cannot be set out or substantiated in detail. However, these are conditions in which it is possible to conduct the analysis, identify potential relationships between constructs and assess the explanatory power of the proposed model.

The model is contextualised for application in NHS hospitals through the analysis of policy documents specific to the hospitals in England. Therefore, the conclusions relating to the influence of the target boundary and gradient on the systemic characteristics are particular to that context and time. The new government has made changes to the policy of targets, which makes it impossible to replicate the context of the study. However, the approach and methods adopted is repeatable.

As noted in Chapter 4, there is a debate about the ability to generalise from case studies. The concern about being able to generalise arises from a particular paradigm view of

research which is criticised (Flyvbjerg, 2006). It is accepted that this research is limited in the number, scope and data gathered about the hospitals investigated. Three NHS hospitals with different characteristics were selected to be studied. The context of their local PCT and SHA were kept constant for CS 1 and 2 to reduce the variation in the external context of the hospitals where original data collection took place. CS 3 relied on secondary sources and therefore does not have the richness of data that is available from the different collection methods used in CS 1 and 2.

The use of system dynamic modelling is limited in that mathematical models are not used to simulate the flows and stocks. It is argued that it is acceptable to use SD models in a qualitative manner to illustrate the flows and create CLDs (Lane, 2008).

There is the limitation of the researcher reflexivity. Prior experience and personal bias can influence the data gathering and analysis (Finlay, 2002, Johnson and Duberley, 2003). The personal career and interests of the researcher are explicitly recognised and acknowledged in the thesis. The triangulation of data and the sharing and testing of thoughts during the research process is a means to counteract such bias.

The timing of the data gathering in CS 1 and 2 during the winter months provided a particular insight into hospitals working during periods of perturbation and continuous stress. The situations investigated do not occur every month. However, the research design enabled rich data to be gathered by a number of means to allow triangulation that increases the reliability of the results. Staff were on the whole very honest and willing to share their views, concerns and ideas.

10.5 Further research

There are a number of issues arising from this study that require further research, which are set out in this section. The issues are arranged into three areas. The first is to investigate in detail the proposition of the five system archetype behaviours, the second is the operational and patient safety matters found during the empirical inquiry and third, the further development of the model and its application.

10.5.1 System Behaviours

The argument is made in Chapter 9 that the interactions of the system construct sets create five emergent types of system behaviour. More detailed research is required into

articulating the markers of such behaviour and the key intervention points to prevent the system increasing the risk to patients. It is argued that the buffer management of the bottleneck resources is a point of intervention that will reduce the danger to patients. The argument needs to be tested.

10.5.2 Operational and patient safety management

There is little evidence of research examining the relationship between the operational management of hospitals and the impact on patient safety. This research provides evidence of the impact of the medical department having to borrow capacity from surgery by outlying medical patients. There is little if any literature on the relationship between medical outliers and the consequences for patient safety. Part of the further work required in this area is to research the impact on patients being moved for non-clinical reasons and being looked after by staff without the expertise for their condition.

There are potential safety implications of moving patients for non-clinical reasons. These include the potential for error occurring with more frequent patient handovers between wards and clinical teams and the increased workload of staff. There is also the need to assess the impact on the clinical care of patients. There is some evidence of delays in treatment but little systematic examination of the consequences of delays.

The batching of medical work, both by GPs and medical staff in hospitals create peaks in the flow of patients through the hospital. There is little research into the consequences for the operational management of the hospital or the safety of patients from such an approach to managing the workload.

The final area of future research in this area is to establish key measures of demand, stocks, flows and discharges for operational managers to use to better anticipate, plan for and adapt to perturbations and periods of continuous stress.

10.5.3 Development of the model

Rasmussen's (1997) SWE is used widely in the literature. This research has developed that model and applied it in a particular context. Research is required to further assess the validity and reliability as a conceptual model in other hospitals in England and potentially more widely in health and other industries where there is a production and safety requirement. The four boundary envelope model is a development of the original

model with potentially wider application than NHS hospitals in England. Research is needed to see if that model provides greater explanatory power than the original one.

Based on the idea from Miller and Xiao (2007) a number of measureable indicators of adaptive or buffer capacity are proposed in Section 9.3.6 as a way to assess resilience. These indicators are drawn from the case studies examined. Data on those specific indicators has not been collected in the detail necessary to assess whether they are valid, reliable and sufficient indicators of resilience as suggested. Different industries will have a variety of other indicators. Research is required to study whether there are specific and or generic indicators of buffer capacity as a means to assess resilience of a system. The use of quantitative SD models could be further investigated as a means to assess the required size of the buffer capacity of the key stocks in the patient admission pathway.

One of the findings from the case studies is that staff at different levels in the hospital responded to the gradient pressures and boundary constraints in a variety of ways. It is suggested that a further development of the SWE is the idea of a hierarchy of interactive envelopes within a system based on the idea of 'panarchy' (Holling, 2001). Exploring this idea is beyond the scope of this thesis but provides potential for future work.

References

- Adger, W. W. (2000) Social and ecological resilience: are they related? *Progress in Human Geography*, 24, 347-364.
- Amalberti, R., Vincent, C., Auroy, Y. & De Saint Maurice, G. (2006) Violations and migrations in health care: a framework for understanding and management. *Qual Saf Health Care*, 15, i66-71.
- Archer, M., Bhaskar, R., Collier, A., Lawson, T. & Norrie, A. (Eds.) (1998) *Critical Realism: Essential Readings*, London, Routledge.
- Argyris, C. (1999) *On Organizational Learning*, Oxford, Blackwells.
- Armitage, G. (2009) Human error theory: relevance to nurse management. *Journal of Nursing Management*, 17, 193-202.
- Armitage, M. & Raza, T. (2002) A consultant physician in acute medicine: the Bournemouth model of managing increasing numbers of medical emergency admissions. *Clin Med JRCPL*, 2, 331-3.
- Ashby, W. R. (1961) *An Introduction to Cybernetics*, London, Chapman & Hall.
- Audit Commission (2009) Auditors' Local Evaluation for NHS Trusts - Guidance for NHS Trusts, Audit 2009/10. London, Audit Commission.
- Bagust, A., Place, M. & Posnett, J. W. (1999) Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *BMJ (British Medical Journal)*, 319, 155-8.
- Barber, M. (2008) *Instruction to Deliver*, London, Methuen.
- Bartunek, J. (2007) Academic-Practitioner Collaboration need not require joint or relevant research: toward a relational scholarship of integration. *Academy of Management Journal*, 50, 1323-1333.
- Bates, D. W., Larizgoitia, I., Prasopa-Plaizier, N., Jha, A. K. & On Behalf of the Research Priority Setting Working Group of the, W. H. O. W. A. F. P. S. (2009) Global priorities for patient safety research. *BMJ*, 338, b1775-.
- Beauchamp, T. L. & Childress, J. F. (1989) *Principles of Biomedical Ethics 3rd Edition*, Oxford, Oxford University Press.
- Bergman, M. M. (2008) The Straw Men of the Qualitative-Quantitative Divide and their influence on Mixed Methods Research. In Bergman, M. M. (Ed.) *Advanced in Mixed Methods Research*. London, Sage.
- Berkes, F. (2007) Understanding uncertainty and reducing vulnerability: lessons from resilience thinking. *Natural Hazards*, 41, 283-295.
- Berwick, D. (1991) Controlling Variation in Health Care: A Consultation from Walter Shewart. *Medical Care*, 29, 1212-1225.
- Bevan, G. (2006) Setting Targets for Health Care Performance. *National Institute Economic Review*, 197, 67-79.
- Bevan, G. (2009) Have targets done more harm than good in the English NHS? No. *BMJ*, 338, a3129.
- Bevan, G. & Hood, C. (2006) What's measured is what matters: targets and gaming in the English public health care system. *Public Administration*, 84, 517-538.
- Bhaskar, R. (1978) *A Realist Theory of Science*, Brighton, Harvester.
- Bhaskar, R. (1979) *The Possibility of Naturalism*, Brighton, Harvester-Wheatsheaf.
- Bhaskar, R. (1989) *Reclaiming Reality: A Critical Introduction to Contemporary Philosophy*, London, Verso.
- Blackstone, J. H. (2004) On the shape of protective capacity in a simple line. *International Journal of Production Research*, 42, 629-637.

- Boaden, R. & Burnes, B. (2009) Health care safety and organisational change. In Hurwitz, B. & Sheikh, A. (Eds.) *Health Care Errors and Patient Safety*. BMJ Books, Chicester.
- Boin, A. & McConnell, A. (2007) Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management*, 15, 50-59.
- Booth Sweeney, L. & Sterman, J. D. (2007) Thinking about systems: student and teacher conceptions of natural and social systems. *System Dynamics Review*, 23, 285-311.
- Boseley, S., Zorlu, G. & Evans, R. (2010) Huge disparity in NHS death rates revealed. *Guardian*. London.
- Bourdieu, P. (1977) *Outline of a Theory of Practice*, Cambridge, CUP.
- Braithwaite, J., Runciman, W. B. & Merry, A. F. (2009) Towards safer, better healthcare: harnessing the natural properties of complex sociotechnical systems. *Qual Saf Health Care*, 18, 37-41.
- Brannick, T. & Coghlan, D. (2006) Reflexivity in Management and Business Research: What Do We Mean? *Irish Journal of Management*, 27, 143-160.
- Brennan, T. A., Leape, L. L., Laird, N. M., Herbert, L., Localio, A. R., Lawthers, A. G., Newhouse, J. P. & Weiler, P. C. (1991) Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study 1. *New England Journal of Medicine*, 324, 370-376.
- Bryman, A. (2004) *Social Research Methods (2nd Ed)*, Oxford, OUP.
- Bryman, A. & Bell, E. (2007) *Business Research Methods (2nd Ed)*, Oxford, OUP.
- Buchanan, D. A., Fitzgerald, L. & Ketley, D. (Eds.) (2007) *The Sustainability and Spread of Organizational Change: Modernizing Healthcare*, London, Routledge.
- Buchanan, D. A. & Storey, J. (2010) Don't stop the clock: manipulating hospital waiting lists. *Journal of Health Organization and Management*, 24, 343-360.
- Burnett, S., Benn, J., Pinto, A., Parand, A., Iskander, S. & Vincent, C. (2010) Organisational readiness: exploring the preconditions for success in organisation-wide patient safety improvement programmes. *Quality and Safety in Health Care*, 10.1136/qshc.2008.030759, -.
- Burrell, G. & Morgan, G. (1979) *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*, Aldershot, Gower.
- Cameron, P. (2006) Hospital overcrowding: a threat to patient safety? *Medical Journal of Australia*, 184, 203-204.
- Carayon, P. (2006) Human factors of complex sociotechnical systems. *Applied Ergonomics*, 37, 525-535.
- Carayon, P., Schoofs Hundt, A., Karsh, B. T., Gurses, A. P., Alvarado, C. J., Smith, M. & Flatley Brennan, P. (2006) Work system design for patient safety: the SEIPS model. *Qual Saf Health Care*, 15, i50-58.
- Care Quality Commission (2009) NHS Performance Ratings 2008/09. CQC, London.
- Care Quality Commission (2010) Performance Assessment 2009/10 - Indicators for acute and specialist trusts. London, CQC.
- Carroll, J. S. & Quijada, M. A. (2004) Redirecting traditional professional values to support safety: changing organisational culture in health care. *Quality and Safety in Health Care*, 13, ii16-ii21.
- Carthey, J., De Leval, M. R. & Reason, J. T. (2001) Institutional resilience in healthcare systems. *Qual Health Care*, 10, 29-32.
- Chang, A., Schyve, P. M., Croteau, R. J., O'leary, D. S. & Loeb, J. M. (2005) The JCAHO patient safety event taxonomy: a standardized terminology and

- classification schema for near misses and adverse events. *Int J Qual Health Care*, 17, 95 - 105.
- Checkland, P. (1981) *Systems Thinking, Systems Practice*, Wiley, Chichester.
- Cilliers, P. (1998) *Complexity and Postmodernism*, London, Routledge.
- Collier, A. (1994) *Critical Realism: An Introduction to Roy Bhaskar's Philosophy*, London, Verso.
- Colwyn Jones, T. & Dugdale, D. (1998) Theory of Constraints: Transforming Ideas? *The British Accounting Review*, 30, 73-91.
- Cook, R. & Rasmussen, J. (2005) "Going solid": a model of system dynamics and consequences for patient safety. *Qual Saf Health Care*, 14, 130-134.
- Cook, R., Woods, D. D. & Miller, C. A. (1998) A tale of Two Stories: Contrasting Views of Patient Safety. Washington, US National Patient Safety Foundation.
- Cook, R. I. & Woods, D. D. (1994) Operating at the Sharp End: The Complexity of Human Error. In Bogner, M. S. (Ed.) *Human Error in Medicine*. Hillsdale, NJ., Lawrence Erlbaum Associates.
- Creswell, J. W. (2007) *Qualitative Inquiry and Research Design*, Thousand Oaks, CA, Sage.
- Creswell, J. W. & Plano Clark, V. L. (2007) *Designing and Conducting Mixed Methods Research*, Thousand Oaks, Sage.
- Crilly, T. & Le Grand, J. (2004) The motivation and behaviour of hospital Trusts. *Social Science & Medicine*, 58, 1809-1823.
- Cropp, R. & Gabric, A. (2002) Ecosystem Adaptation: Do Ecosystems Maximise Resilience? *Ecology*, 83, 2019-2026.
- Cummings, T. G., Neil, J. S. & Paul, B. B. (2001) Closed and Open Systems: Organizational. *International Encyclopedia of the Social & Behavioral Sciences*. Oxford, Pergamon.
- De Vries, E. N., Ramrattan, M. A., Smorenburg, S. M., Gouma, D. J. & Boermeester, M. A. (2008) The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care*, 17, 216-223.
- Dekker, S. (2005) Why we Need New Accident Models. Lund, Lund University School of Aviation.
- Dekker, S. (2006) *The Field Guide to Understanding Human Error*, Aldershot, Ashgate.
- Dekker, S. (2007) *Just Culture: Balancing Safety and Accountability*, Aldershot, Ashgate.
- Dekker, S. (2011) *Drift to Failure*, Farnham, Ashgate.
- Dekker, S., Cilliers, P. & Hofmeyr, J.-H. (2011) The complexity of failure: Implications of complexity theory for safety investigations. *Safety Science*, 49, 939-945.
- Department of Health (2000a) The NHS Plan. London, Stationary Office.
- Department of Health (2000b) An Organisation with a Memory. London, The Stationary Office.
- Department of Health (2001) Building a Safer NHS for Patients: implementing an organisation with a memory. London, Stationary Office.
- Department of Health (2002) NHS Performance Ratings: Acute Trusts, Specialist Trusts, Ambulance Trusts, Mental Health Trusts 2001/02. London, Stationary Office.
- Department of Health (2003) Building the Assurance Framework: a practical guide for NHS Boards. London, Stationary Office.
- Department of Health (2005a) Creating a Patient-led NHS, Delivering the NHS Improvement Plan. London, Stationary Office.

- Department of Health (2005b) Supporting People with Long Term Conditions. An NHS and Social Care Model to support local Innovation and Integration. London, Stationary Office.
- Department of Health (2006a) The NHS in England: the operating framework for 2006/07. London, Stationary Office.
- Department of Health (2006b) The NHS in England: the operating framework for 2007/08. London, Stationary Office.
- Department of Health (2006c) Safety First. London, Stationary Office.
- Department of Health (2007) Operating Framework. For the NHS in England 2008/09. London, Stationary Office.
- Department of Health (2008a) Developing the NHS Performance Regime. London, Stationary Office.
- Department of Health (2008b) High Quality Care for All. London, Stationary Office.
- Department of Health (2008c) Operating Framework for 2009/10 for the NHS in England. London, Stationary Office.
- Department of Health (2009a) The European Working Time Directive - UK notification of Derogation for Doctors in Training. London, Stationary Office.
- Department of Health (2009b) Operating Framework for 2010/11 for the NHS in England. London, Stationary Office.
- Department of Health (2010a) The 18-week rules suite - definitions (Version 5). London, Department of Health.
- Department of Health (2010b) The 18-week rules suite - national clock rules (Version 5). London, Department of Health.
- Department of Health (2010c) Enhanced Recovery. London, Department of Health.
- Department of Health (2010d) Expanding the list of "never events". London, Stationary Office.
- Department of Health (2010e) Quality Accounts Toolkit. London, Stationary Office.
- Diehl, E. & Serman, J. D. (1995) Effects of Feedback Complexity on Dynamic Decision Making. *Organizational Behavior and Human Decision Processes*, 62, 198-215.
- Donabedian, A. (1966) Evaluating the Quality of Medical Care. *The Milbank Memorial Fund Quarterly*, 44, 166-203.
- Donabedian, A. (2003) *An Introduction to Quality Assurance in Health Care*, Oxford, OUP.
- Doyle, J., K. & Ford, D., N. (1998) Mental models concepts for system dynamics research. *System Dynamics Review*, 14, 3-29.
- Eisenhardt, K. M. (1989) Building Theories from Case Study Research. *The Academy of Management Review*, 14, 532-550.
- Fatovich, D. M., Nagree, Y. & Sprivulis, P. (2005) Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia. *Emergency Medicine Journal*, 22, 351-354.
- Finlay, L. (2002) "Outing" the Researcher: The Provenance, Process, and Practice of Reflexivity. *Qualitative Health Research*, 12, 531-545.
- Flick, U. (2006) *An Introduction to Qualitative Research (3rd Ed)*, London, Sage.
- Flick, U. (2007) *Doing Interviews*, London, Sage.
- Flin, R., O'connor, P. & Crichton, M. (2008) *Safety at the Sharp End*, Aldershot, Ashgate.
- Flyvbjerg, B. (2001) *Making social science matter: why social inquiry fails and how it can succeed again*, Cambridge, CUP.
- Flyvbjerg, B. (2006) Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12, 219-245.
- Foot, C. (2011) How do Quality Accounts Measure Up? London, King's Fund.

- Forester, A. J., Clark, H. D., Menard, A. & Et Al (2004) Adverse events affecting medical patients following discharge from hospital. *Canadian Medical Association Journal*, 170, 345-9.
- Forester, A. J., Murff, H. J., Peterson, J. F., Gandhi, T. K. & Bates, D. W. (2003) The incidence and severity of adverse events affecting patients after discharge from hospital. *Annals of Internal Medicine*, 138, 161-7.
- Forrester, J. (1958) Industrial dynamics: a major break-through for decision making. *Harvard Business Review*, 35, 37-66.
- Forrester, J. (1961) *Industrial Dynamics*, Cambridge, MA, MIT Press.
- Forrester, J. (1969) *The Principles of Systems*, Cambridge, MA, Wright-Allen Press.
- Francis Qc, R. (2010a) Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust January 2005 - March 2009 Vol 1. London, Stationary Office.
- Francis Qc, R. (2010b) Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust January 2005 - March 2009 Vol 2. London, Stationary Office.
- Friesdorf, W., Buss, B. & Marsolek, I. (2007) Patient safety by treatment standardization and process navigation - a systems ergonomics management concept. *Theoretical Issues in Ergonomics Science*, 8, 469-479.
- Gardiner, S. C. & Blackstone, J. H. (1998) Dynamic buffering. *International Journal of Production Research*, 36, 333-342.
- Gawande, A. (2010) *The Checklist Manifesto*, London, Profile Books.
- George, A. & Bennett, A. (2005) *Case Studies and Theory Development in the Social Sciences*, MIT Press, Cambridge, Massachusetts.
- Gerring, J. (2007) *Case Study Research, Principles and Practices*, CUP, Cambridge.
- Giddens, A. (1976) *New Rules of the Sociological Method*, London, Hutchinson.
- Giddens, A. (1979) *Central Problems in Social Theory*, London, MacMillan.
- Giddens, A. (1984) *The Constitution of Society*, Oxford, Polity Press.
- Gillham, B. (2000) *Case Study Research Methods*, London, Continium.
- Ginsburg, L. S., Tregunno, D., Norton, P. & Casebeer, A. L. (2006) Patient Safety Culture in Health Organizations: Whose Culture Is It Anyway? In Casebeer, A., Harrison, A. & Mark, A. (Eds.) *Innovations in Health Care*. Basingstoke, Palgrave Macmillan.
- Godlee, F. (2007) Routine reporting. *BMJ*, 334, 0-.
- Goldratt, E. & Cox, J. (2004) *The Goal: A Process of Ongoing Improvement*, Aldershot, Gower.
- Grenier-Sennelier, C., Lombard, I., Jeny-Loeper, C., Maillet-Gouret, M.-C. & Minvielle, E. (2002) Designing adverse event prevention programs using quality management methods: the case of falls in hospital. *Int J Qual Health Care*, 14, 419-426.
- Guba, E. G. & Lincoln, Y. S. (1994) Competing Paradigms in Qualitative Research. In Denzin, N. K. & Lincoln, T. S. (Eds.) *Handbook of Qualitative Research*. Thousand Oaks; CA, Sage.
- Guba, E. G. & Lincoln, Y. S. ((1994)) Competing Paradigms in qualitative research. In Denzin, N. K. & Lincoln, T. S. (Eds.) *Handbook of Qualitative Research*. Thousand Oaks; CA, Sage.
- Gubb, J. (2009) Have targets done more harm than good in the English NHS? Yes. *BMJ*, 338, a3130.
- Hall, L. M., Doran, D. & Pink, G. H. (2004) Nurse Staffing Models, Nursing Hours, and Patient Safety Outcomes. *Journal of Nursing Administration*, 34, 41-45.
- Ham, C. (2009) *Health Policy in Britain*, Palgrave MacMillan, Basingstoke.

- Ham, C., York, N., Sutch, S. & Shaw, R. (2003) Hospital bed utilisation in the NHS, Kaiser Permanente, and the US Medicare programme: analysis of routine data. *BMJ*, 327, 1257-.
- Hammersley, M. (2008) Troubles with Triangulation. In Bergman, M. M. (Ed.) *Advances in Mixed Method Research*. London, Sage.
- Healey, F., Scobie, S., Oliver, D., Pryce, A., Thomson, R. & Glampson, B. (2008) Falls in English and Welsh hospitals: a national observational study based on retrospective analysis of 12 months of patient safety incident reports. *Qual Saf Health Care*, 17, 424-430.
- Health Care Commission (2009) *Safely does It, implementing safer care for patients*. London, Commission for Healthcare Audit and Inspection.
- Healthcare Commission (2006a) *The annual health check: assessing and rating the NHS*. London, Healthcare Commission.
- Healthcare Commission (2006b) *Investigation into outbreaks of Clostridium difficile at Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust*. London, Commission for Healthcare Audit and Inspection.
- Healthcare Commission (2008) *Criteria for assessing core standards in 2008/09 - Acute trusts*. London, Healthcare Commission.
- Healthcare Commission (2009a) *Investigation into Mid Staffordshire NHS Foundation Trust*. London, Healthcare Commission.
- Healthcare Commission (2009b) *Safely does It, implementing safer care for patients*. London, Commission for Healthcare Audit and Inspection.
- Healy, M. & Perry, C. (2000) Comprehensive criteria to judge validity and reliability or qualitative research within the realism paradigm. *Qualitative Market Research: An International Journal*, 3, 118-126.
- Held, D. & Thompson, J. B. (Eds.) (1989) *Social Theory of modern societies; Anthony Giddens and his critics*, Cambridge, CUP.
- Holliday, A. (2002) *Doing and Writing Qualitative Research*, London, Sage.
- Holling, C. S. (1973) Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, 4, 1-23.
- Holling, C. S. (2001) Understanding the Complexity of Economic, Ecological and Social Systems. *Ecosystems*, 4, 390-405.
- Hollnagel, E. (2004) *Barriers and Accident Prevention*, Ashgate, Aldershot.
- Hollnagel, E. (2006) Resilience - the Challenge of the Unstable. In Hollnagel, E., Woods, D. D. & Leveson, N. (Eds.) *Resilience Engineering; Concepts and Precepts*. Aldershot, Ashgate.
- Hollnagel, E. (2009a) *The ETTO Principle: Efficiency-Thoroughness Trade Off*, Ashgate, Aldershot.
- Hollnagel, E. (2009b) The Four Cornerstones of Resilience Engineering. In Nemeth, C., Hollnagel, E. & Dekker, S. (Eds.) *Preparation and Restoration*. Farnham, Ashgate.
- Hollnagel, E., Woods, D. D. & Levenson, N. (Eds.) (2006) *Resilience Engineering; Concepts and Precepts*, Aldershot, Ashgate.
- Hood, C. (2006) Gaming in Targetworld: The Targets Approach to Managing British Public Services. *Public Administration Review*, 66, 515-521.
- Hood, C. & Dixon, R. (2010) The Political Payoff from Performance Target Systems: No-Brainer or No-Gainer? *Journal of Public Administration Research and Theory*, 20, I281-I298.
- House of Commons Health Committee (2006) *NHS Deficits*. London, Stationary Office.
- House of Commons Public Accounts Committee (2009) *Reducing Healthcare Associated Infection in Hospitals in England*. London, The Stationary Office.

- Illich, I. (1997) *Limits to Medicine*, London, Penguin.
- Jackson, M. C. (2003) *Systems Thinking: Creative Holism for Managers*, Chicester, John Wiley & Sons.
- Jeffcott, S. A., Evans, S. M., Cameron, P. A., Chin, G. S. M. & Ibrahim, J. E. (2009a) Improving measurement in clinical handover. *Qual Saf Health Care*, 18, 272-276.
- Jeffcott, S. A., Ibrahim, J. E. & Cameron, P. A. (2009b) Resilience in healthcare and clinical handover. *Qual Saf Health Care*, 18, 256-260.
- Jha, A. K., Prasopa-Plaizier, N., Larizgoitia, I. & Bates, D. W. (2010) Patient safety research: an overview of the global evidence. *Quality and Safety in Health Care*, 19, 42-47.
- Johnson, P. & Duberley, J. (2000) *Understanding Management Research*, London, Sage.
- Johnson, P. & Duberley, J. (2003) Reflexivity in Management Research. *Journal of Management Studies*, 40, 1279-1303.
- Joyce, P., Boaden, R. & Esmail, A. (2005) Managing Risk: A Taxonomy of Error in Health Policy. *Health Care Analysis*, 13, 337-345.
- Karsh, B. T. & Brown, R. (2010) Macroergonomics and patient safety: The impact of levels on theory, measurement, analysis and intervention in patient safety research. *Applied Ergonomics*, 41, 674-681.
- Katz, D. & Kahn, R. L. (1966) *The Social Psychology of Organizations*, New York, John Wiley.
- Kc, D. S. & Terwiesch, C. (2009) Impact of Workload on Service Time and Patient Safety: An Econometric Analysis of Hospital Operations. *Management Science*, 55, 1486-1498.
- Kelman, S. & Friedman, J. N. (2009) Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19, 917-946.
- Kennedy, I. (2001) The Report of the Public Inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-95 *Learning from Bristol*. London, Secretary of State for Health.
- Kernick, D. (Ed.) (2004) *Complexity and Healthcare Organization: A view from the street*, Oxford, Radcliffe Medical Press.
- King, A. (2004) *The Structure of Social Theory*, London, Routledge.
- Klassen, K. J. & Rohleder, T. R. (2001) Combining Operations and Marketing to Manage Capacity and Demand in Services. *The Service Industries Journal*, 21, 1-30.
- Klein, H. K. (2004) Seeking the new and the critical in critical realism: déjà vu? *Information and Organization*, 14, 123-144.
- Klein, R. (1995) *The New Politics of the NHS*, London, Longman.
- Kohn, L. T., Corrigan, J. M. & Donaldson, M. S. (2000) *To Err is Human: Building a Safer Health System*, Washington DC, National Academy Press.
- Kpmg (2009) Learning and implications from Mid Staffordshire NHS Foundation Trust. London, KPMG.
- Kripalani, S., Lefevre, F., Phillips, C. O., Williams, M. V., Basaviah, P. & Baker, D. W. (2007) Deficits in Communication and Information Transfer Between Hospital-Based and Primary Care Physicians: Implications for Patient Safety and Continuity of Care. *JAMA*, 297, 831-841.
- Kuhn, T. (1962) *The Structure of Scientific Revolution*, Chicago, University of Chicago Press.

- Kuhn, T. (1970) *The Structure of Scientific Revolution*, Chicago, University of Chicago Press.
- Kvale, S. (2008) *Doing Interviews*, London, Sage.
- La Porte, T. R. (1996) High Reliability Organizations: Unlikely, Demanding and At Risk. *Journal of Contingencies & Crisis Management*, 4, 60.
- Lane, D. C. (1999) Social theory and system dynamics practice. *European Journal of Operational Research*, 113, 501-527.
- Lane, D. C. (2000a) Should system dynamics be described as 'hard' or 'deterministic' systems approach? *System Research Behavioural Science*, 17.
- Lane, D. C. (2000b) Should System Dynamics be Described as a 'Hard or 'Deterministic' Systems Approach? *Systems Research and Behavioral Science*, 17, 3-22.
- Lane, D. C. (2001a) Rerum cognoscere causas: Part I; How do the ideas of system dynamics relate to traditional social theories and the voluntarism/determinism debate? *System Dynamics Review*, 17, 97-118.
- Lane, D. C. (2001b) Rerum cognoscere causas: Part II; Opportunities generated by the agency/structure debate and suggestions for clarifying the social theoretic position of system dynamics. *System Dynamics Review*, 17, 293-309.
- Lane, D. C. (2008) The emergence and use of diagramming in system dynamics: a critical account. *Systems Research and Behavioral Science*, 25, 3-23.
- Lane, D. C. & Husemann, E. (2008a) Steering without Circe: attending to reinforcing loops in social systems. *System Dynamics Review*, 24, 37-61.
- Lane, D. C. & Husemann, E. (2008b) System dynamics mapping of acute patient flows. *Journal of Operational Research Society*, 59, 213-224.
- Lane, D. C., Monefeldt, C. & Rosenhead, J. V. (2000) Looking in the Wrong Place for Healthcare Improvements: A System Dynamics Study of an Accident and Emergency Department. *The Journal of the Operational Research Society*, 51, 518-531.
- Lawson, M. B. (2001) In praise of slack: Time is of the essence. *Academy of Management Executive*, 15, 125-135.
- Leape, L. (1994) Error in Medicine. *Journal of the American Medical Association*, 272, 1851-1857.
- Leape, L., Berwick, D., Clancy, C., Conway, J., Gluck, P., Guest, J., Lawrence, D., Morath, J., O'leary, D., O'Neill, P., Pinakiewicz, D. & Isaac, T. (2009) Transforming healthcare: a safety imperative. *Quality and Safety in Health Care*, 18, 424-428.
- Leape, L. L. (2002) Reporting of Adverse Events. *New England Journal of Medicine*, 347, 1633-39.
- Lewin, R. & Regine, B. (1999) *The Soul at Work: Unleashing the Power of Complexity Science for Business Success*, London, Orion Business Books.
- Lopez, J. & Potter, G. (Eds.) (2001) *After Postmodernism: An Introduction to Critical Realism*, London, Continuum.
- Loukopoulos, L. D., Dismukes, R. K. & Barshi, I. (2009) *The Multitasking Myth*, Farnham, Ashgate.
- Lowe, C. M. (2006) Accidents waiting to happen: the contribution of latent conditions to patient safety. *Qual Saf Health Care*, 15, i72-75.
- Luthar, S. S., Cicchetti, D. & Becker, B. (2000) The Construct of Resilience: A Critical Evaluation and Guidelines for Future Work. *Child Development*, 71, 543-562.
- Luthar, S. S. & Zelazo, L. B. (2003) Research on Resilience: An Integrative Review. In Luthar, S. S. (Ed.) *Resilience and Vulnerability*. Cambridge, CUP.
- Manyena, S. B. (2006) The concept of resilience revisited. *Disasters*, 30, 433-450.

- Mason, J. (2006) Mixing methods in a qualitatively driven way. *Qualitative Research*, 6, 9-25.
- Masten, A. S. (2007) Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology*, 19, 921-930.
- Mautner, T. (Ed.) (2000) *The Penguin Dictionary of Philosophy*, London, Penguin.
- Mayhew, L. & Smith, D. (2008) Using queuing theory to analyse the government's 4 hour completion target in accident and emergency departments. *Health Care Management Science*, 11, 11-21.
- Mcfadden, K. L., Henagan, S. C. & Gowen Iii, C. R. (2009) The patient safety chain: Transformational leadership's effect on patient safety culture, initiatives, and outcomes. *Journal of Operations Management*, 27, 390-404.
- Mclennan, G. (1995) *Pluralism*, Buckingham, Open University Press.
- Meadows, D. H. (2008) *Thinking in Systems*, White River Junction, Vermont, Chelsea Green Publishing.
- Melnyk, S. A. & Handfield, R. B. (1998) May you live in interesting times...the emergence of theory-driven empirical research. *Journal of Operations Management*, 16, 311-319.
- Meredith, J. (1993) Theory Building through Conceptual Models. *International Journal of Operations and Production Management*, 13, 3-11.
- Meredith, J. (1998) Building operations management theory through case and field research. *Journal of Operations Management*, 16, 441-454.
- Merton, R. K. (1968) *Social Theory and Social Structure*, New York, The Free Press.
- Midgley, G. (2000) *Systemic Intervention: Philosophy, Methodology and Practice*, New York, Kluwer.
- Miles, M. B. & Huberman, A. M. (1994) *Qualitative Data Analysis*, Thousand Oaks, Sage.
- Miller, A. & Xiao, Y. (2007) Multi-level strategies to achieve resilience for an organisation operating at capacity: a case study at a trauma centre. *Cogn Tech Work*, 9, 51-66.
- Mingers, J. (2004) Real-izing information systems: critical realism as an underpinning philosophy for information systems. *Information and Organization*, 14, 87-103.
- Mohr, J., Batalden, P. & Barach, P. (2004) Integrating patient safety into the clinical microsystem. *Qual Saf Health Care*, 13, ii34-38.
- Mohr, J. J., Abelson, H. T. & Barach, P. (2002) Creating Effective Leadership for Improving Patient Safety. *Quality Management in Health Care*, 69-78.
- Monitor (2008) Compliance Framework. London, Monitor.
- Monitor (2009) Management Response to the Internal Audit Report on Lessons Learnt from Mid Staffordshire NHS Foundation Trust. London, Monitor.
- Monod, E. (2004) Einstein, Heisenberg, Kant: methodological distinction and conditions of possibilities. *Information and Organization*, 14, 105-121.
- Mooney, H. (2010) Patient Safety: Hole system approach. *Health Service Journal*, 20 May, 22-23.
- Morath, J. M. & Turnbull, J. E. (2005) *To Do No Harm*, San Francisco, Jossey Bass.
- Nembhard, I. M. & Edmondson, A. C. (2006) Making it safe: the effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior*, 27, 941-966.
- Neuendorf, K. A. (2002) *The Content Analysis Guidebook*, Thousand Oaks, California, Sage.
- Olsen, S., Scobie, S., Chapman, E., Sachs, R., Mckee, M., Vincent, C. & Thomson, R. (2008) What can we learn about patient safety from information sources within an acute hospital: a step on the ladder of integrated risk management? *Qual Saf Health Care*, 17, 209-215.

- Patton, M. Q. (2002) *Qualitative Research and Evaluation Methods 3rd Ed*, Thousand Oaks, Sage.
- Perneger, T. (2005) The Swiss cheese model of safety incidents: are there holes in the metaphor? *BMC Health Services Research*, 5, 71.
- Perrow, C. (1984) *Normal Accidents*, New York, Basic Books.
- Perrow, C. (1994) The Limits of Safety: The Enhancements of a Theory of Accidents. *Journal of Contingencies & Crisis Management*, 2, 212.
- Perrow, C. (1999) Organizing to Reduce the Vulnerabilities of Complexity. *Journal of Contingencies and Crisis Management*, 7, 150-155.
- Perrow, C. (2007) *The Next Catastrophe*, Princeton Princeton University Press.
- Perry, C. (1998) Processes of a case study methodology for postgraduate research in marketing. *European Journal of Marketing*, 32, 785-802.
- Pidgeon, N. (1997) The Limits to Safety? Culture, Politics, Learning and Man-Made Disasters. *Journal of Contingencies and Crisis Management*, 5, 1-14.
- Plsek, P. E. & Greenhalgh, T. (2001) Complexity science: The challenge of complexity in health care. *BMJ*, 323, 625-628.
- Plsek, P. E. & Wilson, T. (2001) Complexity science: Complexity, leadership, and management in healthcare organisations. *BMJ*, 323, 746-749.
- Porpora, D. V. (1998) Four concepts of social structure. In Archer, M., Bhaskar, R., Collier, A., Lawson, T. & Norrie, A. (Eds.) *Critical Realism*. London, Routledge.
- Potter, J. (1996) Discourse analysis and constructionist approaches: theoretical background. In Richardson, J. T. (Ed.) *Handbook of Qualitative Research Methods in Psychology and Social Sciences*. Leicester, BPS Books.
- Prasad, A. (2002) The Contest Over meaning: hermeneutics as an Interpretive Methodology for Understanding Texts. *Organizational Research Methods*, 5, 12-33.
- Prasad, A. & Mir, R. (2002) Digging Deep for meaning: A Critical Hermeneutic Analysis of CEO Letters to Shareholders in the Oil Industry. *Journal of Business Communications*, 39, 92-116.
- Prasad, P. (2005) *Crafting Qualitative Research: Working in the postpositive traditions*, Armonk, New York, M E Sharpe.
- Proudlove, N., Boaden, R. & Jorgensen, J. (2007) Developing bed managers: the why and the how. *Journal of Nursing Management*, 15, 34-42.
- Proudlove, N. C., Gordon, K. & Boaden, R. (2003) Can good bed management solve the overcrowding in accident and emergency departments? *Emerg Med J*, 20, 149-155.
- Radnor, Z. (2008) Muddled, massaging, manoeuvring or manipulated?: A typology of organisational gaming. *International Journal of Productivity and Performance Management*, 57, 316-328.
- Radnor, Z. & Lovell, B. (2003) Success factors for implementation of the balanced scorecard in an NHS multi-agency setting. *International Journal of Health Care Quality*, 16, 99.
- Radnor, Z. & McGuire, M. (2004) Performance management in the public sector: fact or fiction? *International Journal of Productivity and Performance Management*, 53, 245-260.
- Radnor, Z. & Walley, P. (2008) Learning to walk before we try to run: Adapting Lean for the public sector. *Public Money & Management*, 28, 13-20.
- Rasmussen, J. (1997) Risk management in a dynamic society: a modelling problem. *Safety Science*, 27, 183-214.
- Reason, J. (1990) *Human Error*, Cambridge, CUP.

- Reason, J. (1995) Understanding adverse events: human factors. *Qual Health Care*, 4, 80 - 89.
- Reason, J. (1997) *Managing the Risks of Organizational Accidents*, Aldershot, Ashgate.
- Reason, J. (2008) *The Human Contribution*, Farnham, Ashgate.
- Reason, J., Hollnagel, E. & Paries, J. (2006) Revisiting the "Swiss Cheese" Model of Accidents. *Project Safebuild*. European Organization for the Safety of Air Navigation.
- Reiman, T. & Oedewald, P. (2007) Assessment of complex sociotechnical systems - Theoretical issues concerning the use of organizational culture and organizational core task concepts. *Safety Science*, 45, 745-768.
- Riege, A. (2003) Validity and reliability tests in case study research: a literature review with "hands on" applications for each research phase. *Qualitative Market Research*, 6, 75-86.
- Rijpma, J. A. (2001) Complexity, Tight-Coupling and Reliability: Connecting Normal Accident Theory and High Reliability Theory. *Journal of Contingencies and Crisis Management*, 5, 15-23.
- Roberts, K. H. (1990) Some Characteristics of One Type of High Reliability Organization. *Organization Science*, 1, 160-176.
- Roberts, K. H. & Bea, R. (2001) Must Accidents Happen? Lessons from high-reliability organizations. *Academy of Management Executive*, 15, 70-78.
- Rodriguez-Ulloa, R. & Paucar-Caceres, A. (2005) Soft System Dynamics Methodology (SSDM): Combining Soft Systems Methodology (SSM) and System Dynamics (SD). *Systemic Practice and Action Research*, 18, 303-334.
- Roelen, A. L. C., Lin, P. H. & Hale, A. R. (2010) Accident models and organisational factors in air transport: The need for multi-method models. *Safety Science*, In Press, doi:10.1016/j.ssci.2010.01.022.
- Rogers, A. E., Hwang, W.-T., Scott, L. D., Aiken, L. H. & Dinges, D. F. (2004) The Working Hours Of Hospital Staff Nurses And Patient Safety. *Health Aff*, 23, 202-212.
- Rubin, H. J. & Rubin, I. S. (1995) *Qualitative Interviewing: The Art of Hearing Data*, Thousand Oaks, Sage.
- Rynes, S. (2007) Let's create a tipping point: what academics and practitioners can do together. *Academy of Management Journal*, 50, 1046.
- Sagan, S. D. (1993) *The Limits of Safety*, Princeton, Princeton University Press.
- Sagan, S. D. (2004) The Problem of Redundancy Problem: Why More Nuclear Security Forces May Produce Less Nuclear Security. *Risk Analysis*, 24, 935-946.
- Salter, B. (2004) *The New Politics of Medicine*, Basingstoke, Palgrave MacMillan.
- Sari, A. B.-A., Sheldon, T. A., Cracknell, A. & Turnbull, A. (2007a) Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ*, 334, 79-.
- Sari, A. B.-A., Sheldon, T. A., Cracknell, A., Turnbull, A., Dobson, Y., Grant, C., Gray, W. & Richardson, A. (2007b) Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care*, 16, 434-439.
- Schulman, P., Roe, E., Van Eeten, M. & De Bruijne, M. (2004) High Reliability and the Management of Critical Infrastructures. *Journal of Contingencies and Crisis Management*, 12, 14-28.
- Scott, C. (2003) Setting safe nurse staffing levels. London, Royal College of Nursing.
- Scott, J. (2001) Where is Social Structure. In Lopez, J. & Potter, G. (Eds.) *After Postmodernism: An introduction to Critical Realism*. London, Continuum.

- Scott, L. D., Rogers, A. E., Hwang, W.-T. & Zhang, Y. (2006) Effects of Critical Care Nurses' Work Hours on Vigilance and Patients' Safety. *Am J Crit Care*, 15, 30-37.
- Searle, J. R. (1995) *The Construction of Social Reality*, London, Penguin.
- Seddon, J. (2008) *Systems Thinking in the Public Sector*, Axminster, Triarchy Press.
- Senge, P. M. (1990) *The Fifth Discipline: the art and practice of a learning organization*, London, Century Business.
- Sharpe, V. A. & Faden, A. I. (1998) *Medical Harm. Historical, Conceptual and Ethical Dimensions of Iatrogenic Illness*, Cambridge, Cambridge University Press.
- Sheridan, T. B. (2008) Risk, Human Error, and System Resilience: Fundamental Ideas. *Human Factors*, 50, 418-426.
- Silverman, D. (1993) *Interpreting Qualitative Data*, London, Sage.
- Silverman, D. (2005) *Doing Qualitative Research*, London, Sage.
- Simon, H. A. (1957) *Administrative Behaviour; a Study of Decision-Making Processes in Administrative Organizations*, New York, McMillan.
- Sittig, D. F. & Singh, H. (2010) A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Quality & Safety in Health Care*, 19, I68-I74.
- Skyttner, L. (2005) *General Systems Theory*, Singapore, World Scientific Publishing.
- Smart, P. K., Tranfield, D., Deasley, P., Levene, R., Rowe, A. & Corley, J. (2003) Integrating 'lean' and 'high reliability' thinking. *Proc Instn Mech Engrs*, 217 733-739.
- Smith, D. & Fischbacher, M. (2009) The changing nature of risk and risk management: The challenge of borders, uncertainty and resilience. *Risk Management*, 11, 1-12.
- Snook, S. A. (2000) *Friendly Fire*, Princeton Princeton University Press.
- Snowden, D. J. & Boone, M. E. (2007) A Leader's Framework for Decision Making. *Havard Business Review*, November.
- Somers, S. (2009) Measuring Resilience Potential: An Adaptive Strategy for Organizational Crisis Planning. *Journal of Contingencies & Crisis Management*, 17, 12-23.
- Sprivulis, P. C., Da Silva, J.-A., Jacobs, I. G., Frazer, A. & Jelinek, G. (2006) The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments. *Medical Journal of Australia*, 184, 208-212.
- Stacey, R. & Griffin, D. (2005) *A Complexity Perspective on Researching Organisations*, London, Routledge.
- Stake, R. E. (1995) *The Art of Case Study Research*, Thousand Oaks, California, Sage.
- Steel, K., Gertman, P. M., Crescenzi, C. & J, A. (1981) Iatrogenic illness on a general medical service at a university hospital. *New England Journal of Medicine*, 304, 638-642.
- Steier, F. (Ed.) (1991) *Research and Reflexivity*, London, Sage.
- Sterman, J. D. (1989) Misperceptions of feedback in dynamic decisionmaking. *Organizational Behaviour and Human Decision Processes*, 43, 301-335.
- Sterman, J. D. (1994) Learning in and about complex systems. *System Dynamics Review*, 10, 291-330.
- Sterman, J. D. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*, Boston, Irwin McGraw-Hill.
- Sterman, J. D. (2001) System Dynamics Modeling: Tools For Learning in a Complex World. *California Management Review*, 43, 8-25.

- Stock, G. N., Mcfadden, K. L. & Gowen Iii, C. R. (2007) Organizational culture, critical success factors, and the reduction of hospital errors. *International Journal of Production Economics*, 106, 368-392.
- Strauss, A. & Corbin, J. M. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Newbury Park, CA., Sage.
- Stuart, I., Mccutcheon, D., Handfield, R., Mclachlin, R. & Samson, D. (2002) Effective case research in operations management - a process perspective *Journal of Operations Management*, 20, 419-433.
- Sweeney, K. (2006) *Complexity in Primary Care: understanding its value*, Abingdon, Radcliffe Medical Press.
- Sweeney, K. & Griffiths, F. (Eds.) (2002) *Complexity and Healthcare: An Introduction*, Abingdon, Radcliffe Medical Press.
- Sweeney, K. & Williams, M. D. (2010) Safety, systems, complexity, and resilience: What makes organizations safe? In Haxby, E., Hunter, D. & Jaggat, S. (Eds.) *An Introduction to Clinical Governance and Patient Safety*. Oxford, Oxford University Press.
- Tashakkori, A. & Teddlie, C. (Eds.) (2003) *Handbook of mixed methods in social and behavioural research*, Thousand Oaks,, Sage.
- Taylor, T. & Ford, D. N. (2006) Tipping point failure and robustness in single development projects. *System Dynamics Review*, 22, 51-71.
- Tew, P. (2001) Reconsidering Literary Interpretation. In Lopez, J. & Potter, G. (Eds.) *After Postmodernism: An Introduction to Critical Realism*. London, Continuum.
- Thompson, J. B. (1989) The theory of structuration. In Held, D. & Thompson, J. B. (Eds.) *Social theory of modern societies; Anthony Giddens and his critics*. Cambridge, CUP.
- Trzeciak, S. & Rivers, E. P. (2003) Emergency department overcrowding in the United States: an emerging threat to patient safety and public health. *Emerg Med J*, 20, 402-405.
- Tucker, A. L. (2004) The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management*, 22, 151-169.
- Tucker, A. L. & Edmondson, A. C. (2003) Why Hospitals Don't Learn from Failures: Organizational and Psychological Dynamics that Inhibit System Change. *California Management Review*, 45, 55-72.
- Turner, B. & Pidgeon, N. F. (1997) *Man-Made Disasters*, Oxford, Butterworth-Heinemann.
- Ursprung, R., Gray, J. E., Edwards, W. H., Horbar, J. D., Nickerson, J., Plsek, P., Shiono, P. H., Suresh, G. K. & Goldmann, D. A. (2005) Real time patient safety audits: improving safety every day. *Qual Saf Health Care*, 14, 284-289.
- Vaughan, D. (1996) *The Challenger Launch Decision*, Chicago, UCP.
- Vaughan, D. (1999) The dark side of organizations: Mistake, misconduct, and disaster. *Annual Review of Sociology*, 25, 271-305.
- Vincent, C. (2006) *Patient Safety*, Oxford, Churchill Livingstone.
- Vincent, C. (2007) Incident reporting and patient safety. *BMJ*, 334, 51-.
- Vincent, C., Moorthy, K., Sarker, S. K., Chang, A. & Darzi, A. W. (2004) Systems Approaches to Surgical Quality and Safety. *Annals of Surgery*, 239, 475-482.
- Vincent, C., Neale, G. & Woloshynowych, M. (2001) Adverse events in British hospitals: preliminary retrospective record review. *BMJ*, 322, 517-519.
- Vincent, C., Taylor-Adams, S. & Stanhope, N. (1998) Framework for analysing risk and safety in clinical medicine. *BMJ*, 316, 1154-1157.
- Vincent, C. A. (1989) Research into medical accidents: a case of negligence? *British Medical Journal*, 299, 1150-1153.

- Vincent, C. A. (2004) Analysis of clinical incidents: a window on the system not a search for root causes. *Qual Saf Health Care*, 13, 242-243.
- Vincent, C. A. (2010) *Patient Safety*, Chichester, Wiley-Blackwell.
- Voss, C. N., Tsiriktsis, N. & Frohlich, M. (2002) Case research in operations management. *International Journal of Operations and Production Management*, 22, 195-219.
- Wachter, R. M. (2008) *Understanding Patient Safety*, New York, McGraw Hill.
- Wacker, J. G. (1998) A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of Operations Management*, 16, 361-385.
- Walshe, K. & Boaden, R. (Eds.) (2006) *Patient Safety: Research into Practice*, Maidenhead, OUP.
- Walshe, K. & Smith, J. (Eds.) (2006) *Healthcare Management*, Maidenhead, Open University Press.
- Waring, J. J. (2005) Beyond blame: cultural barriers to medical incident reporting. *Social Science & Medicine*, 60, 1927-1935.
- Waring, J. J., Harrison, S. & McDonald, R. (2007) A culture of safety or coping? Ritualistic behaviours in the operating theatre. *Journal of Health Research and Policy*, 12, S1:3-9.
- Wears, R. L. (2004) Organisations and safety in health care. *Quality & Safety in Health Care*, 13, 1-1.
- Wears, R. L. (2005) Keep the celebrations short. *Quality & Safety in Health Care*, 14, 154-154.
- Wears, R. L. & Cook, R. I. (2010) Getting Better at Being Worse. *Annals of Emergency Medicine*, 56, 465-467.
- Wears, R. L., Perry, S. J., Anders, S. & Woods, D. D. (2008) Resilience in the Emergency Department. In Hollnagel, E., Nemeth, C. & Dekker, S. (Eds.) *Remaining Sensitive to the Possibility of Failure*. Ashgate, Aldershot.
- Weick, K. E. (1987) Organisational Culture as a Source of High Reliability. *California Management Review*, 29, 288-299.
- Weick, K. E. & Sutcliffe, K. M. (2001) *Managing the Unexpected*, San Francisco, Jossey Bass.
- Weick, K. E. & Sutcliffe, K. M. (2003) Hospitals as Cultures of Entrapment: A Re-Analysis of the Bristol Royal Infirmary. *California Management Review*, 45, 73-84.
- Weingart, S. N. & Page, D. (2004) Implications for practice: challenges for healthcare leaders in fostering patient safety. *Qual Saf Health Care*, 13, ii52-56.
- Weissman, J. (2007) Hospital Workload and Adverse Events. *Medical Care*, 45, 448-455.
- West, D. (2010) Trusts dodge scrutiny by omitting vital measures. *Health Service Journal*. London, emap.
- West, E. (2000) Organisational sources of safety and danger: sociological contributions to the study of adverse events. *Qual Saf Health Care*, 9, 120-126.
- Wheeler, D. J. (2003) *Making Sense of Data*, Knoxville, Tennessee, SPC Press.
- Williams, M. D. (2008) Designing System Resilience in NHS Hospitals. *Ergonomics Society Improving Patient Safety Conference 16-18 July*. Cambridge.
- Williams, M. D., George, B. & Wade, R. (2009) Rebellion against the 'normal' to improve safety for patients. *16th Annual EurOMA Conference*. Goteborg.
- Williams, M. D., George, B. & Wade, R. (2010) Reducing hospital inpatient falls. In Anderson, M. (Ed.) *Contemporary Ergonomics and Human Factors*. Keale, Taylor Francis.

- Williams, M. D. & Smart, P. A. (2009) System resilience and patient safety during a bed crisis in an NHS hospital in England. *16th Annual EurOMA Conference*. Goteborg.
- Williams, M. D. & Smart, P. A. (2010) Patient Safety: a casualty of target success? *International Journal of Public Sector Management*, 23, 416-430.
- Winch, G. & Derrick, S. (2006) Flexible study processes in a "knotty" system dynamics projects. *Systems Research & Behavioral Science*, 23, 497-507.
- Wolstenholme, E. F. (1999) Qualitative vs Quantitative Modelling: The Evolving Balance. *The Journal of the Operational Research Society*, 50, 422-428.
- Woods, D. D. (2006) Essential Characteristics of Resilience. In Hollnagel, E., Woods, D. D. & Leveson, N. (Eds.) *Resilience Engineering, Concepts and Precepts*. Ashgate, Aldershot.
- Woods, D. D. & Cook, R. (2006) Incidents - markers of resilience or brittleness. In Hollnagel, E., Woods, D. D. & Leveson, N. (Eds.) *Resilience Engineering: concepts and precepts*. Aldershot, Ashgate.
- Woods, D. D. & Cook, R. I. (2002) Nine Steps to Move Forward from Error. *Cognition, Technology & Work*, 4, 137-144.
- Woods, D. D., Dekker, S., Cook, R., Johannesen, L. & Sarter, N. (2010) *Behind Human Error*, Farnham, Ashgate.
- Woods, D. D., Patterson, E. & Cook, R. I. (2007) Behind Human Error: Taming Complexity to Improve Patient Safety. In Carayon, P. (Ed.) *Handbook of Human Factors and Ergonomics in Health Care and Patient Safety*. Mahwah, New Jersey, Lawrence Erlbaum Associates.
- Woods, D. D., Schenk, J. & Allen, T. T. (2009) An Initial Comparison of Selected Models of System Resilience. In Nemeth, C., Hollnagel, E. & Dekker, S. (Eds.) *Preparation and Restoration*. Farnham, Ashgate.
- Woods, D. D. & Wreathrall, J. (2008) Stress-strain Plots as a Basis for Assessing System Resilience. In Hollnagel, E., Nemeth, C. & Dekker, S. (Eds.) *Remaining Sensitive to the Possibility of Failure*. Ashgate, Aldershot.
- Wreathrall, J. (2006) Properties of Resilient Organisations: An Initial View. In Hollnagel, E., Woods, D. D. & Leveson, N. (Eds.) *Resilience Engineering: concepts and precepts*. Ashgate, Aldershot.
- Yin, R. K. (2003) *Case Study Research, Design and Methods*, London, Sage.

Appendices

Appendix 2.1 - Conventions of System Dynamic models

Stock Flow Diagrams (SFDs)

The clouds at either end of a SFD model depict the ‘sink’ or ‘source’ for the flows (Sterman, 2000) (Figure A2.1). A source is where a flow originates from outside the boundary, and a sink is where a flow drains into outside the boundary of the model. Inside the boundary are ‘valves’ which control the rate of flow. There are also information flows which denote policy and decision making. The information flows influence the rate of flow. ‘Stocks’ illustrate the amount of accumulation, namely the difference between inflow and outflow.

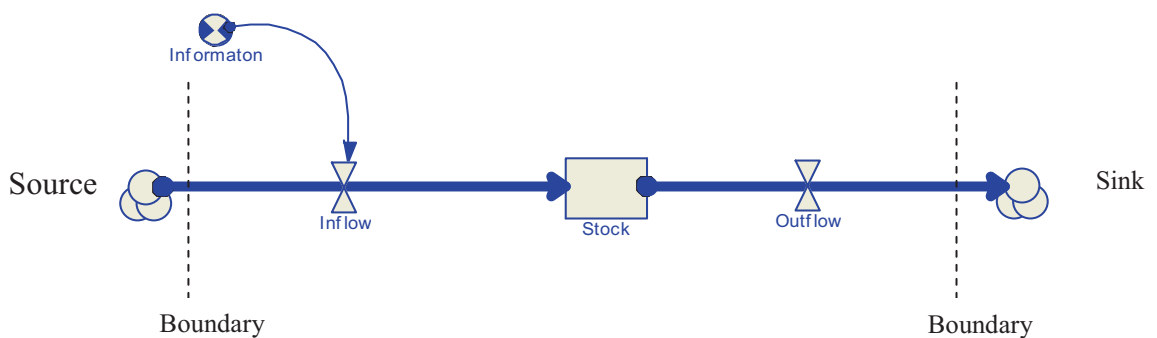


Figure A2.1: Basic structure of a ‘stock flow diagram’

Causal Loop Diagrams (CLDs)

CLDs are used in SD to illustrate the dynamics of the system (feedback loops), (Sterman, 2000, Morecroft, 2007, Sterman, 2001, Senge, 1990). Such an approach is qualitative and cannot be used to build simulation models. The diagrammatical conventions in CLDs are for the variables to be connected by an arrow showing direction of the causal link. For example, the ‘birth rate’ causes the ‘population’ to change (Figure A2.2).

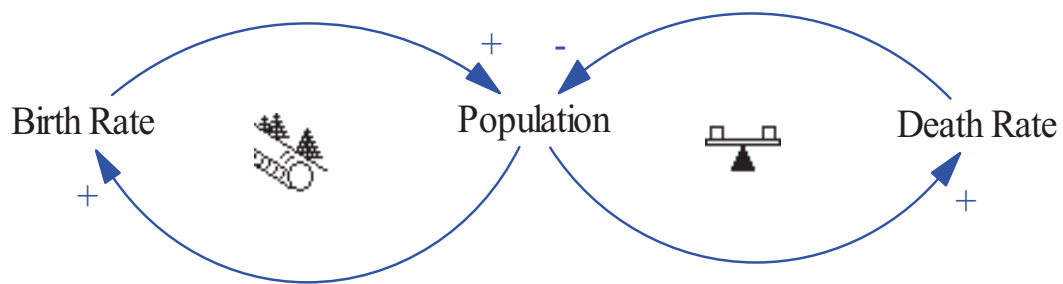


Figure A2.8: Example of a 'casual loop diagram' showing the reinforcing loop generated by the birth rate being 'balanced' by the death rate loop.

Where a '+' or 'S' is used in the notation, the direction of the influence is the same. For example, when the birth rate goes up the population goes up. Where '-' or 'O' is used the direction of influence is the opposite. For example, when the death rate goes up the population goes down.

In CLDs the polarity of the feedback loop is used to illustrate the type of feedback (balancing or reinforcing). Where there is an odd number of – or O, in the loop it has a negative polarity. This is known as a 'balancing' loop. For example, as the 'population' rises the 'death rate' rises. The direction of influence is the same so a '+' sign is used. The loop is balancing because when the 'death rate' rises the 'population' decreases. The direction of influence is the opposite, so a '-' sign is used. Within the loop a logo of a balance indicates that the loop is balancing. A balancing loop means that the system will find a state of equilibrium.

In contrast, where a loop has positive polarity it is known as reinforcing, where exponential growth or decline is experienced. A logo of a snowball gathering speed is used. In large complex systems it is common to find a number of reinforcing loops. These are often counteracted by balancing loops. However, some systems may experience exponential reinforcing growth or decline for a period before the balancing loops take effect.

In modelling systems it is possible to replicate key features of systems such as delay. Delay is another feature of 'stocks' in a system where the outflow is slower than the inflow. Sterman (2000) suggests that perception delays can also occur. This is where decision makers view of the world is influenced by the delay in measurement and reporting on aspects of the process. An example of this could be the perception delay of government policy makers in understanding the dangers to health from smoking.

Figure A2.3 illustrates the features of a balancing loop where there are delays in perception, decision making and action. Decision makers experience a delay in the measurement and reporting of what is actual 'state of the system'. There is also a potential delay due the time taken to use the information to perceive a difference between the actual and desired state of the system. Further delays occur in making decisions and undertaking the necessary corrective action. This can lead to over and under correction. A simple example is a shower with a time delay between changing the temperature control and the water temperature changing. It is common to over correct initially and for the corrections to become smaller over time.

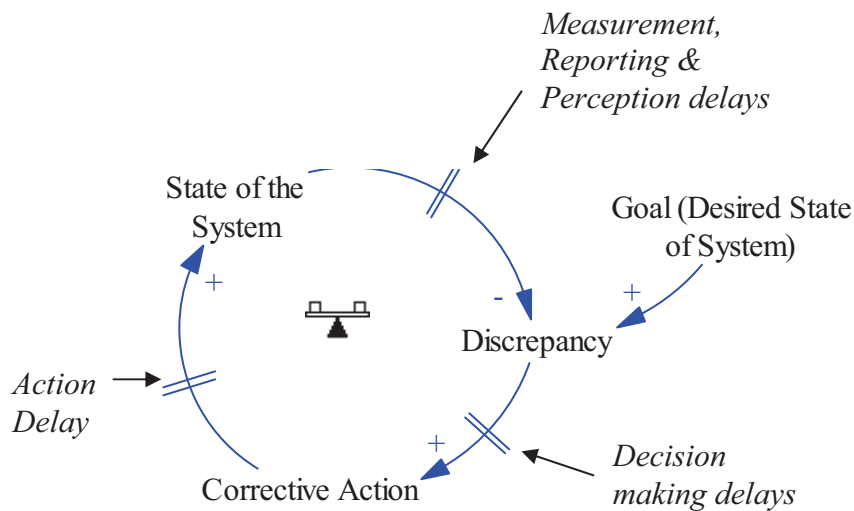


Figure A2.3: Causal loop diagram of oscillating structure and behaviour caused by delays in the balancing loop (adapted from Sterman, 2000)

Appendix 4.1 - List of Documents and Statistical Data Collected

Case Study 1

Doc No	Document Title	Description	Date of Publication	No of Pages
1.1	Annual Report 2007/08	Overview of activity and financial performance	undated	52
1.2	Annual Report 2008/09 and Summary Financial Statements	Overview of activity and financial performance	undated	66
1.3	Escalation Plan	Operational plan to meet high levels of patient demand for beds	9 November 2007	5
1.4	Press Release	Press statement about sickness virus and closure of hospital to visitors	11 December 2008	1
1.5	Annual Control of Infection Report 2008/09	Report on infection outbreaks, compliance, audit and governance	24 June 2009	63
1.6	Daily Bed State Report	Operational report on occupancy, admissions, infections	Daily	2
1.7	Quality Report	First annual review of key quality data	June 2009	15
1.8	Directorate 1 Quarterly Review Q3 2008/09	Performance review data	January 2009	34
1.9	Directorate SS Quarterly Review Q3 2008/09	Performance review data	January 2009	34
1.10				
1.11	Annual Plan 2009/10	Review of performance Future Business Plans	27 May 2009	45
1.12	Our Forward Plans – Summary Strategic Directions 2007-2012	National influences; mission and values; milestones and priorities	undated	14
1.13	Governance Committee	Agenda and Papers for meeting on 3 December 2008	3 December 2008	100 +
1.14	Governance Committee	Minutes of meeting held on 3 December 2008	December 2008	9
1.15	Governance Annual Report 2007/08	Update on achievements and plan for 2008/09	11 September 2008	30
1.16	Governance Committee Decisions Briefing	Briefing to Trust Board on decisions taken July – Sept 2008	28 October 2008	3
1.17	Assurance Framework and risk register	Update to Trust Board	28 October 2008	7
1.18	Healthcare Commission – Annual health check 2007/08	Assessment of ‘Quality of Services’ and ‘Use of Resources’	undated	12
1.19	Performance Reports	Monthly reports to Trust Board 2008 /09	April 08 – March 09	13 - 50
1.20	Matron’s Patient Safety Handover Tool	Presentation slides introducing items on handover list	11 February 2009	7 slides

Case Study 2

Doc No	Document Title	Description	Date of Publication	No of Pages
2.1	Vision, Aims & Values	Board statement	7 July 2009	4
2.2	Annual Report 2008/09 and Accounts	Overview of activity and financial performance	undated	116
2.3	Annual Report 2009/10	Overview of activity and financial performance	undated	53
2.4	Clinical Governance Report 2008/09	Safety, Audit, Governance, Compliance, Patient focus, Environment, Public Health	undated	33
2.5	Clinical Governance Report 2009/10	Safety, Audit, Governance, Compliance, Patient focus, Environment, Public Health	undated	49
2.6	Clinical Governance Committee Minutes	Minutes of Sept and Oct 2009 meetings	3 November 2009	19
2.7	Quality Account 2009/10	Vision, Values, Assurance statement, quality performance	undated	37
2.8	Continuous Improvement Programme Report	Report on 11 areas of improvement including finance, standards for better health, productivity and IT	January 2009	14
2.9	Patient Safety Project Plan	Draft Project Initiation Document	4 February 2009	15
2.10	Clinical Operations Performance Reports	Report to Trust Board on KPIs, Care Quality Standards, MRSA and C.Diff, Standards for Better Health and HSMR	Monthly (Nov 09 – Apr 10)	15
2.11	Infection Prevention & Control Report	Report to Trust Board on C.Diff, MRSA and Hand Hygiene	Monthly (Nov 09 – Apr 10)	5
2.12	Planning for Winter 2009/10	Description of management action to manage any increase in emergency admissions	3 November 2009	16
2.13	Patient Experience Report Q4 2008/09	Patient feedback, complaints and themes	2 June 2009	14
2.14	Finance Report	Financial performance	Monthly (Nov 09 – Apr 10)	7 - 10

Case Study 3

Doc No	Document Title	Description	Date of Publication	No of Pages
3.1	Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust January 2005 – March 2009 Volume 1	Report submitted to the Secretary of State for Health	24 February 2010	455
3.2	Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust January 2005 – March 2009 Volume 2	Report submitted to the Secretary of State for Health	24 February 2010	367
3.3	Opening Statement by Council for Cure the NHS	Statement by legal representative of local pressure group to Inquiry	undated	17
3.4	Opening Statement to Inquiry by Mid Staffordshire NHS Foundation Trust	One page letter from CEO expressing apologies and welcoming Inquiry	undated	1
3.5	Opening Statement by Counsel to the Inquiry	Setting out background and themes of the Inquiry	29 October 2009	18
3.6	Summary of Patient Oral Evidence Week 1	Summary of issues raised by patients or relatives	undated	21
3.7	Summary of Patient Oral Evidence Week 2	Summary of issues raised by 10 patients or relatives	undated	22
3.8	Summary of Patient Oral Evidence Week 3	Summary of issues raised by 12 patients or relatives	undated	30
3.9	Summary of Staff Oral Evidence Week 1	Summary of oral statements from 14 members of staff	undated	22
3.10	Summary of Staff Oral Evidence Week 2	Summary of oral statements from 12 members of staff	undated	29
3.11	Summary of Staff Oral Evidence Week 3	Summary of oral statements from 17 members of staff	undated	32
3.12	Summary of Staff Oral Evidence Week 4	Summary of oral statements from 4 members of staff	undated	11
3.13	Closing Submissions Transcript	Closing statements to the Inquiry	6 January 2010	46
3.14	Joint statement from Monitor and the Care Quality Commission (CQC) on the current position at Mid Staffordshire NHS Foundation Trust		24 February 2010	2
3.15	Healthcare Commission Investigation into Mid Staffordshire NHS Foundation Trust	Investigation report into apparently high mortality rates in patients admitted as emergencies to the Trust	March 2009	176
3.16	Care Quality Commission Mid Staffs 3 month follow up report	Assessment of changes and improvements made at the Trust	17 July 2009	4

3.17	Care Quality Commission Mid Staffs 6 month follow up report	Assessment of changes and improvements made at the Trust	Dec 09	24
3.18	Department of Health: Terms of Reference – Mid Staffordshire NHS Foundation Trust: Lessons Learned Inquiry	Terms of reference for two reviews into lessons learnt	Undated	2
3.19	Prof G Alberti Report	A review of the procedures for emergency admissions and treatment, and progress against the recommendation of the March Healthcare Commission report	29 April 2009	21
3.20	Dr D C Thomé Report	A review of lessons learnt for commissioners and performance managers following the Healthcare Commission investigation	29 April 2009	34
3.21	Government response to Alberti and Colin-Thomé Reports	Summary of findings and actions to be taken at Mid Staffordshire Hospital and the CQC	undated	9
3.22	Written Ministerial Statement – Mid Staffordshire NHS Foundation Trust	Government accepts recommendations of the Alberti and Colin-Thomé Reports	30 April 2009	2
3.23	Mid Staffordshire NHS Foundation Trust: Confidence in Care Transformational Programme – Goals	Report to Trust Board on actions being taken to improve standards of care	17 November 2009	29
3.24	KPMG: Learning and implications from Mid Staffordshire NHS Foundation Trust	Internal Audit Report by KPMG for Monitor establishing lessons to be learnt	5 August 2009	24
3.25	Monitor: Management Response to the Internal Audit Report on Lessons Learnt from Mid Staffordshire NHS Foundation Trust	Response and proposed actions resulting from the Internal Audit Report	3 September 2009	17

Statistical data (Excel Spreadsheets) collected from Case Study 1

xls #	Data File	Description	Size
1.1	Daily bed state AM & PM	Twice daily bed predictor, discharges, empty bed, ED breaches, ITU occupancy, MRSA, C.Diff (Oct 08 – Mar 09)	120 – 134KB
1.2	Length of Stay by Directorate	Average Length of stay by month (April 06 – Feb 09)	36.5KB
1.3	Outliers and plus 15 day length of stay	Number of Outliers & Number of 15 day plus stays (Jan – Oct 09)	418KB
1.4	Quality Dashboard September 2009	Data on HSMR, Adverse events, Incidents, C,Diff, MRSA, Patient experience	204KB
1.5	Bed Occupancy	Bed occupancy by day and ward (2006-2008)	1,628KB
1.6	Same Day Hospital Cancellations	Cancellations by speciality (2006/7, 2007/08, 2008/09)	86KB
1.7	Inpatient and Daycase Pivot Table	Admission data by Directorate (April 06 – Nov 09)	11.3MB
1.8	Discharges	Discharges by Ward and Speciality (Oct 08 - March 09)	368KB
1.9	Medical Outliers Length of Stay	Average LOS by month for medical patients discharged from medical and surgical wards (April 06 – Feb 09)	40KB
1.10	Acute Ward Data	Daily admission, transfer, deaths and discharge data (July 99 – Oct 08)	470KB
1.11	Hourly, daily, variation	Admission and discharges by hour by day of week (2006/07)	19.8MB

Statistical data (Excel Spreadsheets) collected from Case Study 2

xls #	Data File	Description	Size
2.1	Admissions	Medical Admission and Outliers (Dec 08 0 Jan 10)	129KB
2.2	Daily Bed Statistics	Daily spreadsheet of admissions, empty bed, forecast admissions (Apr 04 -Jan 10)	681KB
2.3	Average Length of Stay	ALOS by ward (May 08 – Dec 09)	74KB
2.4	Discharges	Discharges and ALOS by year (07/08 – 09/10)	16,337KB

Appendix 4.2 - Protocol for on site data collection

Context area	Unit of analysis	Questions	Data collection methods
<i>Nature of actions</i>	Directors / hospital wide processes	What types of actions are taken to manage high levels of patient demand? Are they short or long term; planned / unplanned; formal / informal. How does a) finance, b) targets c) staff workload influence the nature of the actions taken?	Interviews Budget reports Hosp admin data Board minutes Bed mgt records Observation Incident reports
	Speciality	As above	Interviews Speciality budget, activity, performance & incident reports
	Team	As above	Interviews Activity & incident reports Observation
<i>Reason for actions</i>	Directors	Why do you take those actions? How would you describe the scale of the problem being faced in terms of running the service, minor – major? Are there some events for which you have planned responses to avoid problems occurring? Who decides/d to take the action? Who do you tell about the action? How did/does a) finance, b) targets c) staffing influence the reason for the actions taken?	Interview Budget reports Performance reports
	Speciality	As above	As above
	Team	As above	As above plus Staff / patient ratio Staff skill mix
<i>Frequency of actions</i>	Directors	How often do you take those actions? How do you define peaks in demand? How often do they occur?	Interview Observation Activity & performance reports
	Speciality	As above	As above
	Team	As above	As above
<i>Outcomes</i>	Directors	What do you consider are/were	Interview

<i>of actions</i>		<p>the outcomes of the actions? How did those outcomes impact on a) patient safety b) finance c) targets d) staff workload? Would you do that action again? What would you do differently next time? Have there been times when the hospital has not been able to cope? Have the frequency of incidents, infections, staff sickness been affected by peaks in demand?</p>	<p>Budget reports Patient activity data Board reports & minutes Bed mgt records Observation Performance Reports Infection control reports</p>
	Speciality	As above	As above
	Team	As above	As above
<i>Boundary setting & monitoring</i>	Directors	<p>How do you define and monitor a) financial success/failure b) target success/failure c) un/acceptable staff workload d) patient safety failure? How do you measure the above? Are there examples of a new 'normal' being established in terms of accepting what was previous unacceptable?</p>	<p>Interview Risk Management Strategy & Policies Assurance Framework Risk Meeting reports & minutes Incident reports Governance Committee & Board Reports</p>
	Speciality	As above	<p>Interview Reports & minutes Incident reports</p>
	Team	As above	<p>Interview Incident reports</p>
<i>Change learning & developing staff</i>		<p>How do you learn from events to which you have to respond? How is that learning recorded / shared? Are procedures / policies / ways of working / structures changed as a result of the learning? Are staff given the opportunity to reflect on incidents / events?</p>	<p>Interview Incident reports Training records Procedures / Policies cited has having changed</p>

Appendix 4.3 - List of Interviewees

Case Study 1

To insure anonymity names are not used. (* denoted those interviewees who were also observed; ^ denotes finding discussed with)

Person ID	Department	Position	Time in post (yrs)	Time in NHS (yrs)	Job Level	No of interviews	Duration (mins)
CS 1.1	Medicine	Nurse	1	13	2	1	42
CS 1.2	Medicine	Nurse	3.5	30	3	1	27
CS 1.3	Medicine	Nurse	3	15	3	1	32
CS 1.4	Board	Director	11	27	1	1	54
CS 1.5	Board	Director	0.5	15	1	1	33
CS 1.6 ^	Board	Director	8	31	1	2	94
CS 1.7	Board	Director	4	4	1	1	60
CS 1.8	Operations	Manager	1.5	11	2	1	48
CS 1.9	Operations	Manager	6	6	3	1	61
CS 1.10	Medicine	Manager	4	18	3	1	31
CS 1.11	Surgery	Manager	13	40	2	1	37
CS 1.12	Medicine	Manager	4.5	7	2	1	51
CS 1.13	Operations	Manager	6	15	2	1	74
CS 1.14 ^	Medicine	Manager	0.3	10	2	2	63
CS 1.15 ^	Board	Director	6	25	1	2	55
CS 1.16	Surgery	Nurse	7	24	2	1	38
CS 1.17* ^	Operations	Nurse	1.5	16	2	1	56
CS 1.18* ^	Medicine	Nurse	5	19	2	1	64
CS 1.19	Medicine	Nurse	3	10	3	1	22
CS 1.20	Medicine	Nurse	8	22	2	1	50
CS 1.21	Medicine	Nurse	2	19	3	1	65
CS 1.22	Medicine	Nurse	0.8	22	2	1	44
CS 1.23	Medicine	Nurse	6	17	3	1	46
CS 1.24	Board	Director	6	37	1	1	61
CS 1.25	Medicine	Doctor	1	18	2	1	60
CS 1.26	Surgery	Doctor	3	35	2	1	25
CS 1.27	Medicine	Doctor	2.5	15	3	1	35
CS 1.28	Medicine	Doctor	13	22	3	1	47
CS 1.29	Surgery	Doctor	5	16	3	1	44
CS 1.30	Medicine	Doctor	2	15	3	1	56
CS 1.31 ^	Medicine	Doctor	0.5	13	3	1	38

In addition to the interviews observations of bed management meetings and processes, the Medical Assessment Unit, the Emergency Department, Senior Nurses, managers, doctors and patient movements, Board, Clinical Risk Committee and management meetings were undertaken over a five month period Oct 08 – Feb 09.

Case Study 2

Access to interviewees was limited to 45 min appointments within which time informed consent was obtained. Some interviewees were either late or had to leave early which curtailed the interview time.

Person ID	Department	Position	Time in post (yrs)	Time in NHS (yrs)	Job Level	No of interviews	Duration (mins)
CS 2.1	Board	Director	5	22	1	1	33
CS 2.2	Board	Director	3	28	1	1	36
CS 2.3	Operations	Manager	10	12	2	1	30
CS 2.4	Medicine	Doctor	1.5	14	3	1	35
CS 2.5	Board	Director	3	14	1	1	29
CS 2.6	Operations	Manager	3	8	2	1	18
CS 2.7	Board	Director	4	26	1	1	34
CS 2.8	Medicine	Manager	3.5	33	2	1	27
CS 2.9	Medicine	Doctor	7	15	1	1	37
CS 2.10	Medicine	Nurse	4.5	30	3	1	31
CS 2.11	Medicine	Nurse	3	23	2	1	34
CS 2.12	Medicine	Doctor	6	18	2	1	30
CS 2.13	Board	Manager	3.5	19	1	1	17
CS 2.14	Medicine	Nurse	2	18	3	1	31
CS 2.15	Medicine	Doctor	1.5	6	3	1	22

In addition to the interviews observations of medical ward rounds, bed management meetings and processes, Medical Assessment Unit and the Emergency Department were undertaken over a three week period Dec 09 – Jan 10.

Appendix 4.4 - Participant Questionnaire

All the information provided on this form will be kept confidential and all results anonymised

Name (please print)

Job Title

e-mail@rdefn.nhs.uk Tel

Time in current post (yrs) Time in NHS (yrs)

1. How do you think the **Trust Board members** rank the following priorities when making decisions: *(Please put in rank order; 1 as highest – 4 as lowest)*

Achieving targets		Adequate staffing		Patient Safety		Achieving financial results	
-------------------	--	-------------------	--	----------------	--	-----------------------------	--

2. How do you think the **Divisional Management teams** rank the following priorities when making decisions: *(Please put in rank order; 1 as highest – 4 as lowest)*

Achieving targets		Adequate staffing		Patient Safety		Achieving financial results	
-------------------	--	-------------------	--	----------------	--	-----------------------------	--

3. How do you think the **clinical teams on wards** rank the following priorities when making decisions: *(Please put in rank order; 1 as highest – 4 as lowest)*

Achieving targets		Adequate staffing		Patient Safety		Achieving financial results	
-------------------	--	-------------------	--	----------------	--	-----------------------------	--

4. How do you think your **line manager** ranks the following priorities when making decisions: *(Please put in rank order; 1 as highest – 4 as lowest)*

Achieving targets		Adequate staffing		Patient Safety		Achieving financial results	
-------------------	--	-------------------	--	----------------	--	-----------------------------	--

5. How would your staff think **you** rank the following priorities when making decisions:

Achieving targets		Adequate staffing		Patient Safety		Achieving financial results	
-------------------	--	-------------------	--	----------------	--	-----------------------------	--

Appendix 4.5 - INFORMATION ABOUT THE RESEARCH

7 October 2008

Title of Project: **System Resilience in NHS Hospitals**

Name of Researcher: Mike D Williams

1.0 Introduction

Your name has been suggested as someone who might be able to contribute to this research project. Before you decide if you wish to participate you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. You may ask me any questions about the study if you wish.

Part 1 tells you about the purpose of the research and what it would involve for you. Part 2 provides more detailed information about the study. Ask us if there is anything that is not clear or if you wish further information. Take time to decide whether or not you wish to take part.

PART 1

1.1 Purpose of the study

The purpose of the study is twofold.

1. To improve the understanding of a particular aspect of patient safety; namely how staff in hospital systems work to avoid things going wrong when under pressure.
2. To fulfil the educational requirements of a PhD.

This will be done by conducting a case study in two NHS hospitals. I will be studying what actions staff take, at various levels in the hospitals, during peaks in demand for beds. The plan is then to develop the theoretical and practical understanding of how such actions constitute resilience within the hospital system and how resilience relates to safety for patients.

1.2 Why you have been invited to participate

Your name has been put forward because you have a particular insight and experience of the area being studied through the job you do in the hospital. You have been chosen after discussions and agreement with the Director with lead responsibility for the study within the Trust.

1.3 Do I have to take part?

It is for you to decide; it is entirely voluntary. I will describe the study and go through this information sheet which will then be given to you. I will then ask you to complete a Consent Form to show that you have agreed to participate. You are free to withdraw at any time without giving a reason.

1.4 What will I need to do if I take part?

If you are asked to be interviewed that will last about 30-45 minutes. Interviews will be conducted at a place and time of your convenience. Permission will be sought to record the interviews to enable the researcher analyse your contribution. You will have the opportunity to see and comment on the transcription of the interviews. If you are willing, you may be asked to comment on any conclusions that are drawn by the researcher.

If you agree to be observed, either in meetings you attend or doing other work during the project duration of three to four months, your continued permission will be sought. You may ask the researcher to leave at any point without giving a reason.

A small number of staff (3-4) will be asked to participate in a short follow up interview (no more than 30 minutes) some six months later.

1.5 Will the information I provide be kept confidential?

The information you provide will be kept confidential and all interview recordings will be destroyed at the end of the study (no later than the end of 2011). All data used in study results will be anonymised.

If this information has interested you then please read Part 2 before making a decision.

PART 2

2.1 What happens if I decide to withdraw?

You may withdraw at any point without giving any reason. Any information collected from you before you withdraw will not be used in the study unless you give permission otherwise.

2.2 What will happen to the results of the study?

The main outcome of the study will be a PhD Thesis which will be stored electronically by the University of Exeter. Academic papers will be published and conference presentations made to share the results.

2.3 Who is organising and funding the study?

The study is being organised by a member of staff (Mike Williams) within the University of Exeter Business School as part of his PhD. The NHS South West is funding the research.

2.4 Who has reviewed this study?

The researcher's supervisor, Professor R Maull, at the University of Exeter Business School has reviewed the study and given it University ethical approval. The Director of Postgraduate Research, Professor I Ng, has undertaken a scientific assessment of the research proposal. Comments made in that report have been accommodated in the current study design.

The NHS Devon and Torbay Research Ethic Committee has also reviewed the study and given their approval.

2.5 Researcher's duty of care to patients

The researcher, as well as NHS staff, has a duty of care to patients. In exceptional circumstances should the research discover issues that raise concerns about the *immediate* safety of patients, the researcher will discuss with you the need to break confidentiality and report the matter to a line manager.

2.6 Further information and contact details

If you wish to contact the researcher for any reason the details are:

Mike Williams 07909 817228 or 01392 262595 / 262557 (University) or
m.d.a.williams@ex.ac.uk

2.7 Complaints

If you wish to raise a complaint or concerns about the research please contact Professor Roger Maull, r.s.maul@exeter.ac.uk or Prof Andi Smart, p.a.smart@exeter.ac.uk or phone 01392 262557

Appendix 4.7 – Ethical Approval



National Research Ethics Service

Our ref LCH

10 October 2008

Mr Mike D A Williams
NHS Research Fellow
University of Exeter
Business School, Streatham Court,
Rennes Drive
Exeter
EX4 4PU

Devon & Torbay REC

Dept of Research Ethics
Old Kenn Ward
Royal Devon & Exeter Hospital (Wonford)
Barrack Road
EXETER
EX2 5DW

Tel: 01392 402369
Fax: 01392 402369

Lesley.Holman@rdefn.nhs.uk

Dear Mr Williams

Full title of study: Developing a mid range theory about 'system resilience' and its relation to patient safety in NHS hospitals in England.

REC reference number: 08/H0202/128

Thank you for your email of 09 October 2008, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

It is noted that you do not wish to detail the safety issues and adverse events in the information sheet as you will seek to obtain the views of staff about this during the project. It is therefore acceptable to ignore this request by the Committee

Thank you also for clarifying that you do not intend to make any changes to the study before it commences in North Devon, other than to amend only local information in the documents.

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA). The favourable opinion for the study applies to all sites involved in the research. There is no requirement for other Local Research Ethics Committees to be informed or SSA to be carried out at each site.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission at NHS sites ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Evidence of Insurance		01 August 2008
Letter of Invitation to Participant - Observations	1.1	31 August 2008
Letter of Invitation to Participant - Interviews	1.1	31 August 2008
Supervisor CV	Roger Maull	31 August 2008
Participant Consent Form	1.1	31 August 2008
Interview Schedules/Topic Guides	8.0 Appendix 1	08 August 2008
Peer Review		07 August 2008
Protocol	8.0	08 August 2008
Investigator CV	Mike Williams	
Application	1.1	31 August 2008
Response to Request for Further Information		09 October 2008
Participant Information Sheet	1.2	07 October 2008

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

08/H0202/128

Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Yours sincerely



KC Miss Kate Caldwell OBE
Chair

Enclosures: "After ethical review – guidance for researchers" SL- AR2

Copy to: Prof Roger S Maull, University of Exeter

Miss Lisa Bower, R&D Department, Noy Scott House, RD&E (W)



National Research Ethics Service

South West 2 Research Ethics Committee

Royal Devon & Exeter Hospital (Heavitree)
Room TH5
Gladstone Road
Exeter
Devon
EX1 2ED

Telephone: 01392 405272

Fax: 01392 405270

Email: Southwest.REC@nhs.net

www.nres.npsa.nhs.uk

18 November 2010

Mr Mike D A Williams
NHS Research Fellow
Business School, Streatham Court,
Rennes Drive
Exeter
EX4 4PU

Dear Mr Williams

Study title: Developing a mid range theory about 'system resilience' and its relation to patient safety in NHS hospitals in England.
REC reference: 08/H0202/128

Thank you for sending the progress report for the above study dated 14 November 2010. The report will be reviewed by the Chair of the Research Ethics Committee, and I will let you know if any further information is requested.

The favourable ethical opinion for the study continues to apply for the duration of the research.

08/H0202/128:	Please quote this number on all correspondence
----------------------	---

Yours sincerely

Mrs Barbara Inger
Committee Co-ordinator

E-mail: Barbara.Inger@nhs.net

Copy to: Prof Roger S Maull, University of Exeter

This Research Ethics Committee is an advisory committee to South West Strategic Health Authority

The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

Appendix 5.1 - Summary of key points from NHS Operating Framework Documents (2006/07 – 2009/10)

Research Memorandum – contents and key points from OF documents

Hermeneutical assumptions: based on experience of working in the NHS, the influence of additional resources since the NHS Plan publication in 2000 meant that the waiting time targets had to be met. The political background was one of significant reform to the structure of the system on the back of the increased investment. In 2005/06 the achievement of targets appeared to take precedence in some organisations to the detriment of financial control. Therefore 2006/07 saw the increasing requirement to get the financial position stabilised.

The NHS in England: the operating framework for 2006/7

Contents

Foreword from Sir Nigel Crisp.....	1
1. Introduction and context.....	2
2. Priorities for 2006/7.....	5
3. Choice and commissioning in 2006/7.....	9
4. Providing NHS services in 2006/7.....	12
5. Payment by Results in 2006/7.....	17
6. System management and regulation.....	21
Annexes	
A. Financial management in 2006/7.....	24
B. National targets and Local Delivery Plan (LDP) measures.....	25
C. Demand management in 2006/7.....	28
D. Interim SHA system management role in 2006/7.....	30
E. Transition arrangements: January to July 2006.....	31

The introduction speaks of moving the NHS from a top down target led service to one that has incentives to respond to patients. Annex B lists 34 targets although many others are listed in another document referred to. The main waiting time targets are highlighted are cancer waits and 18 week maximum wait. MRSA reduction target is also emphasised.

The major focus of the document is on financial management and the structural reform.

Nationally, we shall be putting a particular focus in 2006/7 on:

- achieving robust **financial health**;
- pushing forward the implementation of **reform**; and
- achieving **six specific service priorities** derived from the Planning and Priorities Framework.

Six Priorities are:

- Health inequalities:** to deliver the LDP trajectories that make the most progress in reducing health inequalities by 10% by 2010, focusing on life expectancy at birth. The initial focus will be on smoking cessation. We will establish systems for implementation and to track progress for 2007/8 on this and other key interventions, particularly in the spearhead PCTs.

Cancer 31-day and 62-day waits: to ensure the sustained delivery throughout 2006/7 of a maximum waiting time of two months from urgent referral to treatment, and of one month from diagnosis to treatment, for all cancers.

18-week maximum wait: to ensure that by 2008 no one waits more than 18 weeks from GP referral to hospital treatment.

MRSA: to achieve year-on-year reductions in MRSA levels, as set out in the agreed LDPs for 2006/7.

Patient choice and booking: to ensure that every hospital appointment will be booked for the convenience of the patient (by implementing the Choose and Book system) and that every patient is offered a choice of at least four providers.

Sexual health and access to Genito-Urinary Medicine (GUM) clinics: to deliver the 2006/7 LDP trajectories so that by 2008 everyone referred to a GUM clinic should be able to have an appointment within 48 hours.

Apart from financial management structural reform dominates and is summarised in this table:

Implementing reform: expectations of change by March 2007		
	by March 2006	by March 2007
Practice-based commissioning	20% of practices	Universal coverage
Number of PCTs	303	120 to 160+ depending on consultation
Choice of hospital	4+	Extended
Choose and Book	25%	90%
NHS Foundation Trusts	32 (acute)	65 to 80 including 5 to 10 mental health
Independent Sector Treatment Centre (ISTC) capacity	18	24
Payment by Results	£9 billion of services covered	£22 billion of services covered
More service delivered in the community	The forthcoming White Paper will create new levers and incentives for shifting care	

The NHS in England: the operating framework for 2007/08

Contents

Foreword from David Nicholson, NHS Chief Executive	2
1 Introduction and context	4
2 Health and service priorities	8
Development priorities for 2007/08	8
Recovery action to improve mental health services	11
Local action in preparation for 2008/09	12
3 Reforms	14
Patient choice and voice	14
Commissioning	17
Strengthening PCTs	22
Strengthening providers	23
Payment by results	24
Information Management and Technology	26
4 Moving to a modern financial regime	28
Financial regime	29
Application of resource accounting and budgeting to NHS trusts	29
Replacement of cash brokerage with loans	30
Capital finance	30
Strategic reserves	30
Efficiency improvements	31
Central budgets	31
SHA bundle	31
Early notification of financial planning assumptions	32
Annex A Timetable	33
Annex B National Targets and Local Delivery Plan Requirements	35
Annex C Principles for the NHS	39
Annex D Roles, responsibilities and business processes (including special circumstances) for 2007/08	42

New CEO for the NHS – Sir Nigel Crisp lost his job due to poor financial management. The new CEO sets out his view in the Foreword to include:

“In order to ensure focus and consistency of purpose, there are no new national priorities for the service to deliver on for this year. But the operating framework for 2007/08 does set out a number of changes in the way we expect the NHS to conduct its business that are significant. We are deliberately moving towards a much more rules based system which will bring some much needed rigour and transparency to the NHS.

...The operating framework for 2007/08 provides consistency of purpose for the NHS, setting out the key targets that our staff need to focus on in order to improve patient experience, reduce health inequalities and achieve financial health.

...The operating framework sets out the importance of tackling all healthcare associated infections, and instead of setting a new national target requires PCTs to engage with clinicians and to agree local targets for reducing levels of *Clostridium difficile*. Achieving 18 weeks is a national target, but we recognise that the degree of transformation required cannot be delivered from the centre, and that it is local clinicians and managers who need to drive this change.”

2007/08 is stated as being the second stage of system reform following the NHS Plan (2000). Given the financial problems it is not surprising that a major focus of this document is on tightening the financial management processes. Makes no mention of the workload of staff. Safety is mentioned x6; linked to the wider phrase of quality, safety, etc. The focus on MRSA is linked to the need to reduce avoidable suffering deaths.

There are a series of actions to be taken – the impression given is that take all the actions and major improvements in the system will occur and the patients will benefit.

Development priorities for 2007/08

Four issues will require particular attention by all organisations in 2007/08 because of both the degree of challenge they pose and their importance to public confidence in the NHS. These are:

- achieving a maximum wait of 18 weeks from GP referral to start of treatment
- of patients;
- reducing rates of MRSA and other healthcare associated infections;
- reducing health inequalities and promoting health and well-being;
- achieving financial health.

Annex B sets out the many targets similar to previous year but now with delivery dates set out. The real emphasis is on the achieving the 18 week target and infection control.

Annex C sets out the Principles of the NHS – which includes #3 ‘We will work continuously to improve quality and safety’. It mentions issues such as learning and reducing mistakes and complying with national inspection and regulation.

#8 ‘We will support and value our staff’ – education, training and development.

The operating framework. For the NHS in England 2008/09

Contents

Foreword 2

1 Overall context 4

2 Priorities 8

Overall approach 8

National priorities for 2008/09 10

Cleanliness and healthcare-associated infections 11

Improving access 12

Keeping adults and children well, improving their health and reducing health inequalities 15

Experience, satisfaction and engagement 18

Emergency preparedness 19

Local priorities 20

Recovery and preparatory action 20

3 Enabling strategies 24

Empowering patients – choice, information and personalisation 24

World-class commissioning 26

Practice-based commissioning 27

Specialised services commissioning 27

National contract 28

Commissioning services for military personnel and veterans 30

System management 30

Provider development 31

Leadership and workforce 33

Information 35

4 Financial regime 38

PCT allocations 39

Resource accounting and budgeting 40

Financially challenged trusts 40

Payment by Results 41

Efficiency 42

SHA bundle 43

Central budgets 43

Information for financial planning 43

5 Business processes 44

Contract implementation 45

Roles and responsibilities 45

The NHS's contribution to Local Area Agreements 48

Timetable 49

Annex A: Existing commitments 50

Annex B: Planning process timetable 52

Annex C: System management timetable of products 53

Annex D: Principles and rules for cooperation and competition 55

Annex E: Contract 56

Priorities:

- improving cleanliness and reducing healthcare-associated infections;

- improving access through achieving 18-week referral to treatment and better access to GP and primary care services;
- keeping people well, improving overall health and reducing health inequalities;
- ensuring we improve the patient experience, staff satisfaction and engagement; and
- not being found wanting in our preparations to respond to emergencies such as an outbreak of pandemic flu.

Section on enabling strategies which, as well as system reform, mentions workforce and the need for leadership and engagement. After a period of expansion benefits now need to be realised.

Safety is mentioned once in relation to specialist commissioning. The emphasis on the targets to reduce MRSA and C.Diff is present as a means to lessen the harm to patients. Targets are a major feature with Annex A setting out more clearly than previous years the waiting time targets.

Annex A

Whilst there is a need to focus on new priorities, it is essential that the levels of service set through previous commitments, which should have been achieved by April 2008, are maintained. We will ask the Healthcare Commission to feed the following specific commitments into its performance assessment of NHS bodies, alongside its performance assessment of other issues:

- four-hour maximum wait in A&E from arrival to admission, transfer or discharge;
- guaranteed access to a primary care professional within 24 hours and to a primary care doctor within 48 hours;
- a maximum wait of 13 weeks for an outpatient appointment;
- a maximum wait of 26 weeks for an inpatient appointment;
- a three-month maximum wait for revascularisation;
- a maximum two-week wait standard for Rapid Access Chest Pain Clinics;
- thrombolysis ‘call to needle’ of at least 68 per cent within 60 minutes, where thrombolysis is the preferred local treatment for heart attack;¹³
- guaranteed access to a genito-urinary medicine clinic within 48 hours of contacting a service;
- all patients who have operations cancelled for non-clinical reasons to be offered another binding date within 28 days, or the patient’s treatment to be funded at the time and hospital of the patient’s choice;
- delayed transfers of care to be maintained at a minimal level;
- all ambulance trusts to respond to 75 per cent of Category A calls within 8 minutes;
- all ambulance trusts to respond to 95 per cent of Category A calls within 19 minutes;
- all ambulance trusts to respond to 95 per cent of Category B calls within 19 minutes;
- a two-week maximum wait from urgent GP referral to first outpatient appointment for all urgent suspected cancer referrals;
- a maximum waiting time of one month from diagnosis to treatment for all cancers;

The operating framework. For the NHS in England 2009/10

Contents

Foreword

Implementing High Quality Care for all	3
1 High Quality Care for all maintaining the momentum	7
2 High Quality Care for Patients and the Public Setting the priorities for 2009/10	11
3 Enabling High Quality Care throughout the NHS a system designed to deliver quality	23
4 Investing in Quality financial framework for 2009/10	33
5 Making Quality Happen Planning and partnership working	41
Annex A: Planning Process timetable	45
Annex B: Existing Commitments	46
Annex C: Vital Signs	47
Annex D: Contract	48

The content page is very different – each heading is related to quality. This reflects the paper ‘High Quality Care for All’ which makes quality as an organising principle for the NHS. Complete contrast to previous year where safety did not feature other than in relation to infection control.

Foreword sets out no new targets but an emphasis on reducing waiting times, tackling infections and financial stability. The NHS had made a surplus in 2008/09 but large scale value for money effort is needed in the light of wider economic difficulties.

The challenge for 2009/10

This Operating Framework therefore sets out a huge leadership challenge, as we are asking the clinical and managerial community to do four things simultaneously:

- Continue to deliver on the national priorities that matter most to our patients and public, so that our progress in these important areas is sustained and improved.
- Invest the additional resources wisely in order to prepare for the need to make substantial efficiency savings in 2010/11 and for a tighter financial climate thereafter.
- Start to put in place the strategic enablers and foundations that will help deliver the ten SHA regional visions and put quality at the heart of all that we do.
- Develop new ways of working and leading that reflect the evidence base and principles for driving large-scale transformational change.

Set within the challenges the national priorities are not changed.

Our five national priorities for 2009/10 were established through the last Operating Framework and remain:

- improving cleanliness and reducing HCAs;
- improving access through achievement of the 18week referral to treatment pledge, and improving access (including at evenings and weekends) to GP services;
- keeping adults and children well, improving their health and reducing health inequalities;

- improving patient experience, staff satisfaction and engagement; and
- preparing to respond in a state of emergency, such as an outbreak of pandemic influenza.

There is a strong approach to quality which has been defined as being:

Defining quality: *High Quality Care for All* set out our ambition for a system wide focus on quality, by setting out a definition of quality covering three specific domains:

- **Safety**—the first dimension of quality must be that we do no harm to patients. This means ensuring that the environment is safe and clean and tackling issues such as healthcare associated infections, where we have made great progress over the past year.
- **Effectiveness**—this includes clinical outcomes, such as mortality and survival rates. Another important aspect of effectiveness is avoiding ill health and helping people to stay healthy. But just as important is the effectiveness of care from a patient perspective, measured through patient reported outcome measures (PROMs).
- **Patient experience**— this includes the quality of care and the delivery of personalised care, focusing on the compassion, dignity and respect with which patients are treated and how easy it is for patients to access services, taking account of the need to promote equality for minority groups. (p.23)

Interestingly, the safety domain focuses on infection and does not consider wider system issues. Effectiveness includes the clinical aspect but this can be seen as how effective clinical practice is, rather than on how the system of healthcare can be effective. The word 'safety occurs x8 in the document.

'Never events' are mentioned (p.43) in terms of reporting, management and payment regimes. Root Cause Analysis are required if a Never event occurs. It appears that the big stick approach will ensure that sufficient barriers are put in place to prevent the events happening.

Leadership of the workforce is again emphasised (p.29) 3% cash releasing savings required (p.37) Workforce planning is encouraged (p.42-3)

Annex C provides a large number of 'Vital Signs' some of which are linked to national targets. Print is so small to fit it on one page it has to be at 125% to read on screen.

Summary: This OF is very different in content and style. Although the key issues on targets and finance are there they are set within a wider context of quality.

Overall the OFs stress the system reforms; financial control; targets and control of MRSA and C.Diff. Although the workforce is mentioned in terms of planning and development, there is no mention of workload on staff or wider safety implications of the need for increased productivity. There is a shift over time as the financial position is stabilised, the reform agenda delivered and most of the targets delivered. The Department moved to a tiering system of targets – small number of national requirements and then more locally agreed performance standards built into contracts.

Appendix 5.2 - Results of the content analysis of the NHS Operating Framework documents

NHS Operating Framework - Content Analysis

	2006-07	2007-08	2008-09	2009-10
Financial Failure				
Financial	28	58	34	33
Finance	8	9	3	15
Savings	0	0	3	6
Budget	4	25	20	7
Cost	8	9	9	4
Overspend	1	2	2	0
Expenditure	0	7	6	2
Income	3	8	10	6
Pay / payment	18	15	21	18
Activity	10	17	23	9
Demand	8	2	4	3
Tariff	20	12	15	29
Surplus	8	7	14	24
TOTAL	116	171	164	156
Mean	8.92	13.15	12.62	12.00
Target Failure				
Target	17	28	22	18
achievement	0	1	3	5
standards	8	18	11	18
compliance	2	3	5	5
wait / waiting time	17	16	14	17
fail / failure	0	0	3	3
access	9	17	43	35
health check	1	0	0	0
TOTAL	54	83	101	101
Mean	6.75	10.38	12.63	12.63

2006-07 2007-08 2008-09 2009-10

Unacceptable Workload

Staff / staffing workload	2	16	36	29
workforce	0	0	0	0
working hours	1	10	16	36
human resources	0	0	0	0
admissions	0	0	0	1
stress / sickness	0	1	2	2
activity	0	0	0	0
occupancy	0	0	0	0
morale	0	0	0	0
turnover	0	0	0	0
training	0	0	0	0
education	1	1	6	7
	0	1	1	8
TOTAL	4	29	61	83
Mean	0.31	2.23	4.69	6.38

Safety Failure

Patient safety / safety	3	6	1	7
Infection / MRSA / C.Diff	6	23	19	21
Incidents	0	0	0	0
Quality	14	30	25	115
Harm / error	0	1	1	2
Medicines mgt / Presc error	0	0	0	0
Adverse events	0	0	0	0
Morbidity	0	0	2	1
Mortality	11	12	2	5
Deteriorating patient	0	0	0	0
TOTAL	34	72	50	151
Mean	3.4	7.2	5	15.1

Appendix 6.1 - Extracts from Field Notebooks 1 - 4

FN 1.24	Bed Mt: 2 pts in Ortho ready for transfer to community hospital. Site Mgr requested they be found place to be 'sat out' whilst waiting for transport. Those 2 beds immediately allocated to ED pts. No question of length of wait for transport or suitability of patients to be sat out.
FN 1.32	"To cancel a patient is the ultimate failure." Lead for Patient Flow
FN 1.42	On call Manager: "Patient safety is paramount - never put an infected patient on a ward area - never over ride infection control. If pt needed to breach ED target that is what happens"
FN 1.44	Site Mgt Office: Sometimes have 3 moves to create a bed on EMU; Surgery – we tell the wards to get the patient ready, send them to theatre. They may have to spend an extra hour in recovery. Recovery are used to that now - previously they would have objected but not now. We communicate with the wards (patients may start on another ward) and Recovery. The anaesthetists get upset when they can't find their patients
FN 1.48	Targets monitored daily - 1 breach cannot be recovered; Finance x 2-3 per month – problem can be recovered; Staffing more difficult but bids are made for more staff (e.g. 8 in endoscopy); Patient safety - every 3 months review of incidents although the emphasise need to keep patient safe.
FN 1.54	Staffing: showed me staffing spreadsheet on computer and talked about her methods of monitoring staff workloads and morale
FN 1.58	Trust Board Meeting - Financial Report: £9M surplus planned = + 2,55M over; Income - moved by £1M but exp rise by £1,7M.
FN 1.74	Governance Committee: Item 8.6 Norovirus outbreak report - no comments made.
FN 1.78	"You can't knit nurses or doctors." Vacancy last winter of 150 WTE, now down to small # - still have sickness problems - more robust HR reports.
FN 2.13	Lead nurses using patient safety handover sheet and brief on staffing to on-call senior nurse
FN 2.23	Empty beds on Gynae Wd; Infection Control require only surgical pts to be outlied as same doctors visit Maternity. Surgical pt less risk of NV.
FN 2.28	Falls: intentional rounding - now 5 falls a fortnights was over 17 over 5 wards. Attention paid to falls - Ward x had two fatalities resulting from falls -

	leaders and staff committed to do something.
FN 2.29	Reflection: Example of deviance of falls moving from being seen as 'normal' to 'not normal'.
FN 2.31	3pm extra bed meeting chaired by COO. 71 outliers; high medical take, thin medical staffing; surgery face problems if ward xx remains medical
FN 2.32	Pre-planned surgery pts to move to Gynae ward
FN 2.34	Debate about types of medical patients that could be moved to Gynae ward: 13 empty beds on Gynae; take 51 = 10 more than expected, therefore need to use Gyane ward. Surg has no pts available to move, There may be med pts who could move to Gynae. But those pts have had several moves. If we need to use those beds we may have to move again to survive. Review again at 4pm bed meeting.
FN 2.34	Nurse AA phoned medical unit co-ordinator to check staffing. Staffing gone to agency - nothing available. 'Have Bank done text to all?' Needing to consider going outside contract for agency staff, Need slack in the system, therefore, have to go to "KKKK" for x 2 every shift. We won't get the quality but we can put them in areas and move others. Asked COO; permission given.
FN 2.36	Gynae Ward - use bay for medical pts. Medical Director advice "avoid elderly and those on antibiotics" - laughter 'there are no young patients'
FN 2.38	After 4pm meeting went with Nurse B to: EMU; Ward B; ED; EMU; Gynae Wd; EMU; Gynae Wd; EMU; ED; her office; Gynae Wd; bed meeting.
FN 2.38	Follow Nurse B to Office - collect bag + drug - go to Gyane Ward and advised Staff Nurse which Medical Wards would stock other drugs needed by medical patients.
FN 2.38	7pm bed meeting: Whiteboard system showing 10 ED breaches
FN 2.39	On EMU the intensity and speed of work was incredible - twice one SHO had Medical Notes but the patient had moved.
FN 2.42	Day case unit keep 1 trolley / bed for inpatient list admission; Patients wait in day room. YY Ward: 5 pts to go into 3 beds; will use day case bed x3 this day; 1 patient for this day case bed being admitted in examination room.
FN 2.48	11.10am ED - 1 pt breached from majors waiting for relative; 1 pt breached from minors needing SAU bed. Surgical pt allocated to bed with med pt in ti waiting for medical review.
FN 2.50	"Sit patients in Ward Y day room" (Site Mgr)
FN 2.52	"Can you do that admission in the office?" (Site Mgr)
FN 2.52	Recovery: Pt in recovery - Ward X request discharge from Recovery; the

	staff nurse had not discharged before.
FN 2.56	ED 15.45hrs: 4 ambulance crews waiting to handover; 6 patients in ED corridor; 10 patients waiting for beds.
FN 2.56	Ambulance crews now held in ED - system gone solid. 'Speed up discharges and be ruthless on who sits out.'
FN 2.56	Ward C - clearing equipment from bay to free up 3 beds; CEO and Director of Nursing going to help; 1645hrs 'can we help make beds' (CEO and DoN)
FN 2.58	Dir Mgr on phone to Ward Sister trying to get patient onto ward (into treatment room) from ED. But already sorted by Bed Mgr.
FN 2.60	Site Management Office: CEO watching site mgrs; CEO spots ED pt with 3 mins to breach - site mgr phones ED. Assured x2 on route to EMU. Phones ED Majors - need pushers so CEO and Director of Nursing leave to do it.
FN 2.64	ED Breach of Four Hour Target Reports from overnight: 5/6 January 2009 00.53hrs – moved patient from Ward H to Ward O. Moved EMU side room to Ward H side room. EMU side room needed cleaning. When S/R ready on EMU, ED explained that patient needed 1:1 nursing. EMU could not give 1:1 nursing; patient went to ITU” 02.16hrs had to transfer patient to side room on Ward L from EMU. EMU side room had to be cleaned before patient could be admitted.” 02.56hrs Patient vomiting. Unexplained cause. Needed side room on EMU. Had to move patient out of side room on Ward T. Transfer patient from EMU. Clean EMU side room before patient could be admitted.”
FN 2.64	Site Mgt Office: Ward XX not willing to take cancer surgery patients." Site Mgr "cancer take priority - send to theatre and hold in Recovery until a bed comes up."
FN 2.66	10.50hrs: Lead Nse for I/C "Who made decision and when about putting medical outliers on Gynae? Now got medical outliers from Ward C, EMU and other areas. You will live to regret that decision."
FN 3.10	Junior medical staffing - running with 8 vacancies, indeed 1 x SHO to 15 patients; reality 1 x SHO per ward plus F1 = 2 Jn Drs for 28 pts. System has no resilience due to lack of med staff and accommodation to be flexed to meet rise in demand.
FN 3.24	No standardised system for developing 'admission list'
FN 4.6	F2 runs take - generates list in his pocket; have to keep phoning to find out if patient seen. Trailing a joint piece of paper.

FN 4.8	11.10hrs Main round - see patient from late take (not overnight); Nurse says 'Gp working practice - gets very busy at 4pm'
FN 4.8	Pt with cardiac 24 hr tape; due off at 4pm. Needs to be analysed by 5pm. If OK then home. Consultant said may have to wait for analysis until tomorrow. Nurse pushed for discharge as no beds
FN 4.8	Patient with history of 1 fit - MRI - outlie. Nse to doctor 'Happy to outlie' Dr to Nurse 'offered enough resistance' Nse to Dr 'you ain't seen nothing yet' No medical beds therefore outlie to Surgery.
FN 4.12	Sun 27th high take 30 - low discharges as not many senior doctors as holiday period. Now backing up
FN 4.12	Bed meeting: debate about patient with fit waiting for MRI - could they not be sat out rather than go to orthopaedics as an outlier? Why not escalated to duty manager?
FN 4.14	Bed meeting: Outliers reviewed individually - plan for care / treatment / discharge. Emphasis - how to speed up discharge - "get on top of outliers; we will have no chance over New Year." "Now we have 21 outliers we have lost the plot." Dir Mgr
FN 4.18	ED full; 2 pts in resus inappropriately; extra nurse in majors; trolleys in corridor x 5; medical expected (GP admissions) in waiting room.
FN 4.18	Dr to see GP admission pt in Paediatric room - moved child and parents into adult minors
FN 4.18	Sister triages all ambulance cases
FN 4.19	Same pattern as MAU and wider hospital over shorter time period = 4 hrs
FN 4.20	Sister went into resus to find bed space for ambulance trolley pt. Not needed as medical patient just got from waiting room to bay 1 has gone to MAU, therefore avoiding breaches.

Appendix 6.2 – Extracts from coded data

Case Study 1

Coded at CS 1: SWE; Patient Safety

[<Internals\CS1 Interviews\ 28; Doctor>](#) - § 6 references coded [9.95% Coverage]

Reference 1 - 2.10% Coverage

Yeah, when we are under pressure the areas where patient safety is definitely vulnerable and examples are we don't have the time or the resources to do adequate observations or monitoring of the patient and therefore we sometimes do not recognise the severity of their complaint as early as we could have done.

Reference 2 - 0.91% Coverage

Clearly as in any other branch of medicine there have been instances of us committing drug errors, which in my view are sometimes but not always ascribable to the pressure of work and this desire to go faster and the lack of time to cross check what we are doing.

Reference 6 - 3.44% Coverage

We have unfortunately had two patients who during the periods when we have been very busy we haven't been able to send a second nurse to triage and on one occasion a patient waited for it was just under an hour, this turned out it was a child who had meningitis and then there was an adult who within a few days who waited a similar period with serious septicaemia and as a consequence of that we have reviewed our policy in terms of when we double team triage and we will send a doctor to triage now or in fact we have been known to send a consultant to triage if we've got a spare consultant. We sometimes have two consultants on during the day and the advantage of certainly a consultant who as they do it (a) hopefully accurately (b) very, very quickly.

[<Internals\CS 1 Interviews\ 29; Doctor>](#) - § 6 references coded [10.00% Coverage]

Reference 2 - 0.60% Coverage

I would expect is to see an increase in error rate and I think the harder people will have to work the more you will side step various safety aspects that are in place

Reference 3 - 1.32% Coverage

I think if you accept the assumption that an anaesthetist going to see the patient before the operation getting to know them and taking their history is designed to improve the safety of the anaesthetic conduct doing that assessment in inadequate environment under time pressure will encourage things to be missed.

[<Internals\CS 1 Interviews\ 22 Nurse>](#) - § 18 references coded [22.75% Coverage]

Reference 5 - 1.48% Coverage

there's also the issue of what happens when somebody is sick on top of your normal staffing level or somebody is very confused and need so you know for each new thing that comes along the system is having to change and adapt to meet a changing need because a person falling over is at high risk of somebody who is very physically unwell and previously I wouldn't say we had that emphasis you know before we had a

confused patient walking around and you know they fall over and that's sort of at the point they fall over whereas now we are having to plan for caring for them.

Reference 6 - 0.63% Coverage

(Referring to outlying patients) It's not good, it's not good. The safest places for the patient to be in the right place with the right doctors looking after them so yes it is a risk but it has to be an evaluated risk.

Reference 9 - 1.03% Coverage

Safety is not just the patients in the building it's those coming through the doors and we have to create capacity for those coming through the doors. By 8 o'clock this morning we did not have a bed empty anywhere

Reference 11 - 1.11% Coverage

...everyday you would be making decisions to do something, cut a corner or do something slightly differently or I haven't got time to do that so that's the bit I am not going to do.

Reference 13 - 3.83% Coverage

I think there are, well IV antibiotics would be an example. In any sort of given medical ward you could have two patients on IV antibiotics with the same level of staffing but just as easily you could have 20 patients on IV antibiotics and each IV antibiotic has a set time it has to be given over, so for every patient that's sort of five, ten, fifteen, twenty minutes' worth of time, to cut the corner is that something that should have been given over four minutes with a nurse sat there giving it slowly is only given over two because (a) somebody is making the decision actually it's safe and not causing, because apparently you won't be but those nurses will probably have the knowledge that policy says they should be given that over four minutes.

Reference 15 - 0.37% Coverage

on the ward next to us they have a large day case list and those patients are prepped, cared for, clerked in the corridor until a bed comes up.

Reference 16 - 0.73% Coverage

There's not you know a suitable waiting area, there's not a day case area for them so there's a lot of care in the corridor, they don't specifically go you know come back from their procedure to the corridor that doesn't happen but they might go from the corridor to their procedure.

[Internals\CS 1 Interviews\ 04 Manager](#) - § 5 references coded [18.10% Coverage]

Reference 3 - 7.07% Coverage

...you are always dealing with risks because actually even if you said we'll close the doors it doesn't actually change the risks the patients here are facing and it probably increases the risk for patients sitting outside on the ambulances. ...I don't remember ever doing anything that we are knowingly, we knowingly do something that is more risky than it should be but we do do things that you wouldn't choose to do if you didn't have to.

Reference 4 - 1.35% Coverage

we had a series of falls and not just, falls where patients come to harm certainly on Ward K and Ward B and probably over a six to eight month period quite significant, quite a significant number so we started doing route cause analysis on all of them and really what came out of that is in terms of that they went through the governance Board and there's a new programme in place on Kenn and Bovey where they do

specialist round in terms of the falls assessment which actually, so those sorts of things come through to the Board members and actually there's a particular process that we are, we look at

[<Internals\CS 1 Interviews\ 19 Nurse>](#) - § 10 references coded [25.15% Coverage]

Reference 2 - 0.87% Coverage

the care that's given is good as long as they've got the right staffing levels. At times I would say when we really do drop down to low numbers due to sickness I would say that probably patients might I wouldn't say they would suffer but they might not have as optimum care as we would like.

Reference 3 - 3.43% Coverage

I would say that staff are so overwhelmed with the number of patients they are actually not clearly focusing on what they need to be focusing on especially if it's junior members of staff and I think it's things like drug administration gets given late because they are worrying about getting patients to theatre so some of the routine things might just go aside and it's not noticed straightaway because the patients would probably not instantly show any signs of suffering or of lack of care but actually over time if you continually don't give drugs on time it might have a knock on effect to those patients but it's almost an unseen risk

Reference 5 - 3.23% Coverage

There was the early warning score which is when you know the patient is deteriorating and we were finding that they weren't either being done or if they were being done it was, they were going to the F1, most junior doctor and they weren't getting past that.

Reference 9 - 2.10% Coverage

are there any things that you monitor on a regular basis in terms of where there's mortality or morbidity?

Well we have all of that but it's, I don't know it doesn't sort of yet, somebody could tell me about all those figures but actually it's, it's, I think all the work that we are doing is influencing a lot of that but I wouldn't be able to say well actually you know I think that we've made our mortality figures go right down I think that's got to be a Trust wide, but I think all these little things that we are doing and a lot of it I think is very good work I have, you know I think will have an impact on the ultimate sort of Trust wide figure.

Reference 10 - 4.43% Coverage

Yeah but that's given if you think about the amount of time I would say that we, say we had an hour for a quarterly review we would spend quarter of an hour, twenty minutes to half an hour talking about the finances and about how we are meeting targets or we did and then the actual bits would be at the end like this last time we really didn't spend much time talking about that (Patient Safety) because there were some other issues around you know particular consultants so when I sat back and sort of listened because that's what I tend to do there is I would say that end of the, at the end of the meeting the amount of effort that goes into discussing that is quite minimal I would say.

[<Internals\CS 1 Interviews\ 18 Nurse](#) - § 3 references coded [4.47% Coverage]

Reference 1 - 0.37% Coverage

as soon as you outlie a patient you begin to build in an inefficiency structure and you

end up with another handover and another handover and things break down you just don't get that level of continuity.

Reference 2 - 1.56% Coverage

I think skill mix is for me a good indicator for patient safety

[<Internals\CS 1 Interviews\ 31 Doctor](#) - § 9 references coded [23.56% Coverage]

Reference 2 - 0.85% Coverage

It's difficult if a patient can't actually get in the door such is the log jam then that actually does seem to jeopardise patient safety and I certainly agree with patients being left in corridors that jeopardises safety but in spite of the quite severe pressure I don't actually see that happening.

Reference 3 - 2.90% Coverage

what about the medical outliers? What's your view as to their safety?

I don't, I don't know if you could demonstrate in any trial that they did less well than the people on a general medical ward yet it's a pain having to go out and look after them and it creates more work but we are professional people so you just do the work in a professional way as it's the same body of work. ...but it doesn't make that much difference so long as you adequately staff these places I think it's all right.

Reference 4 - 4.41% Coverage

how do you measure whether you are keeping patients safe?

I guess you could measure your calamity rate in terms of you get some unexpected deaths in hospital

Reference 6 - 6.36% Coverage

The problems we get then is if we get a kind of unidentified outlier at that time and that's a risk because if we don't know about him or her they can be forgotten about especially on a surgical ward and then you get the call that, you literally get a call sometimes at half past three/four o'clock will you come and see this patient on you know gynae is the classic ward for us but ...oh you've got a medical patient who hasn't been seen post take and you go oh Christ why are you calling me now half past three?

Reference 9 - 0.29% Coverage

I guess that it's your lone wolf patient on the surgical ward that is an issue in this Trust I guess.

[<Internals\CS 1 Interviews\ 7 Director](#) - § 7 references coded [7.19% Coverage]

are there other regular reports that you get related to patient safety?

We get the normal performance reports against the targets that we've got essentially against those targets we get adverse events at the Governance and that's the Board as appropriate. We get the claims against the Board and if there are any exceptional issues that the Medical Director will bring to us if there's anything unusual happened in the hospital and the Medical Director will report to us when those have occurred.

Reference 7 - 0.47% Coverage

if the operation isn't done properly is there a longer recovery period or as a

consequence of that from the technical activity of the operation itself. I am not sure that I see anything which tells me whether that's good or bad at this hospital.

[Internals\CS 1 Interviews\ 6 Director](#) - § 2 references coded [9.05% Coverage]

Reference 1 - 0.67% Coverage

so safety becomes more of an issue and staff I suppose we expect staff to overwork at this point in order to keep the patients as safe as possible

Reference 2 - 8.38% Coverage

We are, we are learning I think like most healthcare organisations of what is acceptable and if you use MRSA as the prime example of that where I don't know three or four years ago MRSA was just a part of being in hospital and nobody could do anything about it and it was just you know you might have MRSA you might survive it or it might kill you but that was just about being in hospital and I think the attitude's changed hugely to our tolerance of MRSA and whether it's an acceptable consequence of healthcare. Similarly with C-Dif you know we were conscious that C-Dif was around but actually was that something that was manageable so those high profile healthcare acquired infection experiences I think have given us a real sort of hang on do we have to tolerate system failure in any of it and are we, how safe are we as an organisation, how do we compare to other organisations not least in health and so we've got a real sense of we want to reduce avoidable patient safety issues wherever we can.

[Internals\CS 1 Interviews\ 9 Manager](#) - § 2 references coded [2.23% Coverage]

Reference 1 - 1.03% Coverage

They had this quarterly review which is from the executive group and they get grilled about the financial status of the directorate every quarter and the lead nurse for theatres said, how long in these quarterly reviews do you spend looking at talking about risk issues and M B said, oh about 10 minutes. And A M, Assistant Directorate Manager said, that's interesting; I've never been asked anything.

Reference 2 - 1.20% Coverage

it's a vicious circle being busy, competing targets, demands, next set of issues and sometimes safety type issues slip down the agenda I feel.

[Internals\CS1 Interviews\ 25 Doctor](#) - § 13 references coded [12.13% Coverage]

Reference 2 - 0.30% Coverage

There are no decent systems to ensure patient safety.

Reference 5 - 0.77% Coverage

So if you said to somebody what is the level of patient safety in this hospital now and what is the individual performance of any member of staff with regard to patient safety nobody could give you that figure. They could give you clinical incident forms, they could give you governance issues but those are all systems that first of all tend to under report and secondly tend to only flag up the extremes

Reference 7 - 0.53% Coverage

so a very, very common safety issue is that medication is delayed because the nurses don't get there you know if you want an audit of quality and timing of the drug round time from prescription to delivery of the drug would be an audit for me of safety if you collected that data

Reference 9 - 1.67% Coverage

I've had something very simple like, right prime example of patient safety we have no system in this hospital for knowing for every emergency medical patient whether they have been seen by a consultant or not. So if we have 70 people admitted one day that spreads to the far corners of the hospital we currently have no idea, no system for identifying where those patients are. What we rely on is usually a junior doctor being around who has stuck a name on a piece of paper.

[<Internals\CS 1 Interviews\ 24 Doctor>](#) - § 6 references coded [4.12% Coverage]

Reference 2 - 1.18% Coverage

In the daytime you know again it's just purely a matter of volume and it can sometimes get too much so that people wait too long and I think the other thing that happens is that the nurses, the nurses get too busy so you have delays which actually we know can harm patients so you know in sepsis we know that every hour you delay starting the antibiotic carries a significant mortality so you know nurses just not having the time to get round to starting things probably and of course it follows on from the doctors delaying in seeing the patient really and giving a diagnosis. So I think getting the right manpower to the people quickly enough

Reference 3 - 0.14% Coverage

you are always making the best of a bad job every time you outlie a patient.

[<Memos\Memo 021109 Bed management observations Dec 08; Jan 09 crisis point>](#) - § 1 reference coded [1.04% Coverage]

Reference 1 - 1.04% Coverage

Manager potentially moving the safety boundary to avoid a breach in ED target by asking for patient to be admitted to a ward treatment room.

Coded at CS 1: SWE; Targets

[<Internals\CS 1 Interviews\ 04 Director>](#) - § 2 references coded [2.21% Coverage]

Reference 2 - 1.11% Coverage

I think as a Board you said you weren't sure whether you were now at fifteen weeks or thirteen weeks RTT but you were assuming it was thirteen. Yes. **So you were cracking on with that as it were.** Yes that's the fact and we can do that because essentially it's only a cash issue because every patient that's thirteen weeks will become fifteen weeks two weeks later so they pay us two weeks late. **Right.** So what. It's only, it's a cash flow debate it's not a patient treatment debate.

[<Internals\CS 1 Interviews\ 18 Nurse>](#) - § 3 references coded [1.60% Coverage]

Reference 1 - 0.19% Coverage

that's why they opened up "L" (daycase) because of the four hour element and there was just nowhere else to go.

Reference 3 - 0.89% Coverage

I know where we are at for all of the targets for the division. I suppose in my position I need to know. I know we are out with the healthcare standards, I know we are out with our Directorate Audit Healthcare standards because I am the one that's pulled it

together I know that we are compliant. I know where we are at with yeah with all of them but I need to know where we are at with all of them.

[<Internals\CS 1 Interviews\ 31 Doctor](#) - § 4 references coded [8.16% Coverage]

Reference 1 - 2.06% Coverage

at Christmas time we were teetering on the very edge of kind of target failure which for this Trust is the kind of ultimate a most heinous crime if you like so it was something that they would be very keen obviously not to step over and in fact you know at times at that time we stepped over the boundary of target failure.

So which target are you thinking of?

My particular area of course if the four hour A&E wait which is, I am part of that because I help to relieve that by getting rid of patients from the ward and thereby being able to accommodate them on the ward and also I can actually physically eject patients directly from ED or make sure they are appropriately seen to make sure that that doesn't happen.

Reference 3 - 1.18% Coverage

what kind of pressure gets applied to you and your colleagues to discharge patients?

We try and discharge patients early but it doesn't actually trigger a kind of switch in my brain that says oh yeah I am going out today with a kind of decreased threshold to sending patients home do you see what I mean. The same criteria still apply, we are guided by I guess our clinical skill and also accepted practice.

Reference 4 - 2.27% Coverage

I think, I've come here from two non-foundation Trusts and the difference, the big difference I guess is the hard nosedness of bed management I guess here in that, I don't think for a minute they would jeopardise patient safety but you know the target is everything and it is rigorously enforced if you like. ... You know it seems a very financially motivated Trust everyone is very much more financially aware as clinicians and kind of aware of targets, aware of you know what it is to be a Foundation Trust and what it means

[<Internals\CS 1 Interviews\ 06 Director](#) - § 2 references coded [6.94% Coverage]

Reference 1 - 3.31% Coverage

As an organisation I think, the complexity of targets is something that I think is extremely difficult in the health service so I would say that the organisation knows the headline targets so the waiting and access targets. Some of the underpinning targets I think are shrouded in mystery to, not only to them but to me at times so the 18 week is a real good example of that where we are achieving the 18 week target but we are failing on data completeness because the algorithm that they are using is a flawed algorithm. So you know you can fail something even though you are doing the right things.

Reference 2 - 3.63% Coverage

Yes. Yes and we do that as far as the divisions are concerned as well through the quarter reporting process. So you know the specific targets that they have to be concerned with they are very clear about them. They also have an understanding of everybody else's the whole complex picture around targets as do the Board, as do the Council of Governors because in the performance report we give them an analysis of the Monitor's targets, the sort of over arching Department of Health targets and the

locally agreed targets so that the three phase performance report that's actually quite clear for the Board and Council of Governors about what those three things are.

[Internals\CS 1 Interviews\ 08 Manager](#) - § 2 references coded [5.22% Coverage]

Reference 1 - 0.59% Coverage

Finance is less of an issue for us, targets are an issue but we are managing that and you know where it does tend to pop over is the whole issue of kind of over working staff or not having quite enough staff and spread things a little bit thinly at certain points in time.

Reference 2 - 4.63% Coverage

Well again the Emergency Department position is highlighted every day and is part of this operational forecast you know and on a weekly basis, well on a daily basis we are looking at where we are against the kind of you know the position for that week so very much saying we need to be achieving 98% every week here. So that's, and this forecast is circulated widely across the organisation every day. At our 12 o'clock bed meeting we review all breaches for the previous day and try and identify any themes so that's managed very much on a daily basis. Our, the referral to treatment target is, we have a referral to treatment steering group and again the, we have information that goes around on a weekly basis to the clinical divisions you know saying where they are against that. ...Others, things that are kind of on the other side of that, I guess within here as well we also talk every day about our CI so actually the C Dif infections so those, so we look at new cases that have come up within the last 24 hours and whether they are on the cohort ward or not and if they are not on the cohort ward what's our plan going to be to move patients back to the right place and all of that is then summarised within our performance report which goes to the Performance Management Group and to the Trust Board monthly.

[Internals\CS 1 Interviews\ 11 Manager](#) - § 2 references coded [2.94% Coverage]

Reference 1 - 2.49% Coverage

I know, well the target boundaries are usually set in the date and the targets that we are, 18 weeks, 15 weeks, four hour you know they are all very clearly marked so we know on a day to day basis where we are with the four hour A&E wait. I know on a weekly basis where I am with the 18 week referral to treatment for admitted and non admitted because that comes to us and on a weekly basis how many patients are on our waiting lists and we know roughly how much it takes to do that level of work so on a very basic level we know whether the waiting listby consultant. So the target, those targets are very clearly defined

[Internals\CS 1 Interviews\ 12 Manager](#) - § 1 reference coded [1.36% Coverage]

Reference 1 - 1.36% Coverage

Even more tightly managed so I think, I think it would be fair to say that targets are the most closely managed because you only have to have one or two kind of failures with a target and you can't take them back whereas with your finances if you overspend in month one you can claw it back over the rest of the year. If you have a breach of some description it sits there. So targets are monitored all the time literally every day I will be looking at a range of target indicators.

[Internals\CS 1 Interviews\ 26 Doctor](#) - § 2 references coded [3.19% Coverage]

Reference 2 - 0.90% Coverage

I actually welcome targets I am probably one of the few doctors who welcomes targets

because I think it has forced the medical profession to change not just the medical profession the whole system to change the way we work,

[<Internals\RDE Interviews\ 17 Nurse>](#) - § 1 reference coded [1.49% Coverage]

Reference 1 - 1.49% Coverage

I mean from our point of view within the operations directorate it's very easy for us you know the analyst, "RB" etc are able to give us that information at the drop of a hat they are always reporting in on our meetings to say how we are doing and the challenge for me is obviously around the four hour target particularly. Monitoring it quarterly you know we have our weekly target that we are always trying to beat 98% but we have our quarterly target that's from an organisation perspective and it's about communicating that to people because I don't think they always realise and we have our own internal targets for instance our own targets to have no more than 120 fifteen day plus stayers.

[<Internals\CS 1 Interviews\ 05 Director](#) - § 1 reference coded [1.59% Coverage]

Reference 1 - 1.59% Coverage

Yes, yes it's monitored monthly by the Board but you know sort of weekly and daily by other key members, divisional managers, Director of Ops, the Chief Executive, myself you know we are sort of keeping an eye on what's happening throughout the month and then the Board formally monitors it on a monthly basis.

[<Internals\CS 1 Interviews\ 24 Doctor](#) - § 1 reference coded [0.86% Coverage]

Reference 1 - 0.86% Coverage

Well that's interesting some of them not too difficult you know four hour wait and the 18 week RTT was quite a challenge to people though and I think that's why some of the early achievers actually didn't because they were they didn't understand the question. RTT was a lot more complex and the stages of treatment a lot more complex than people realised at first I think. I feel pretty confident now.

Coded at CS 1: SWE; Finance

[<Internals\CS 1 Interviews\29 Doctor](#) - § 1 reference coded [0.38% Coverage]

Reference 1 - 0.38% Coverage

I think the organisation probably knows very well I think one of the things they do, do well is you know manage financially very stringently. I have no idea.

[<Internals\CS 1 Interviews\ 22 Nurse](#) - § 3 references coded [2.07% Coverage]

Reference 1 - 0.27% Coverage

Money, not been mentioned, we spend what we've got to so staffing we are trying to get anything we can.

Reference 2 - 1.26% Coverage

It's never mentioned. We do whatever we've got to do, I've got two agency, you no question of it's expensive or it's you know a resource we shouldn't be using we've gone through the normal, there's an escalation try our own staff, try the nurse bank, we will go to agency and we will go sooner rather than later so that we can try and get anything that's available.

Reference 3 - 0.54% Coverage

It is an area that we are trying to plan for in the long term, we know that we are repeatedly opening these areas but in the short term it's, we spend what we've got to make sure we've got staff to provide care

[<Internals\CS 1 Interviews\ 04 Director>](#) - § 7 references coded [5.45% Coverage]

Reference 1 - 1.10% Coverage

So if we were looking at the hierarchy of what was going those two the financial failure and target failure you would put in place before you would want to overwork the staff or go through that boundary or that boundary. Now I think there's a question in terms of when we are as busy as we are you are by default overworking the staff and there's no option other than that. You know you can use that financial situation, you try and alleviate that but in the short term it's not possible.

Reference 2 - 0.72% Coverage

No that's because we've got a big surplus it's very unusual, I'm hiding money at the moment

Reference 3 - 1.68% Coverage

I think, I think we are in a different, we are in an unusual situation because everybody knows we are significantly ahead of plan and therefore the consequences of relapsing aren't as important for us as other organisations or people who aren't in that financial situation that we are in. So I think people know that but they know that's going to be in the context of that they will be held to account at the next quarter review for what they've done. So they won't be, so they are going to have to be able to justify the shifts that have occurred but actually they can justify it in terms of pressure we've been under and it's about maintaining quality and trying to do something about overworking the staff, that's going to be an acceptable reason.

Reference 4 - 0.45% Coverage

And because they know that we're, at a senior level we know what's going on, they don't have the fear that it's going to come as a surprise to us that we've suddenly shifted the financial boundaries.

Reference 5 - 0.56% Coverage

So we've spent a lot of time and quite a lot of resource over the last couple of years just getting the infrastructure and the kit available to a level so that people aren't chasing round looking for pumps, looking for the basic kit that they need.

Reference 7 - 0.34% Coverage

so we invested £5 million last year in terms of revenue, very much targeted at the clinical teams coming up with the things that would make life easier

[<Internals\RDE Interviews\ 18 Nurse>](#) - § 1 reference coded [1.08% Coverage]

Reference 1 - 1.08% Coverage

No because I mean you've spent X amount of agency in a couple of months' time we we'll say actually that's because we had to open up CDU for a two week period and on top of that we had better open [Ward] "C", on top of CDU for a two week period and we had to use agency to, so I can go back and say the spend 'cos they'll say why have you got so much agency – but you know it's not, I am not restricted in doing that it's actually what would just make us safe.

[<Internals\CS 1 Interviews\ 23 Nurse>](#) - § 1 reference coded [0.17% Coverage]

Reference 1 - 0.17% Coverage

I've got a budget to stick to and you know I have to work within in my budget

[<Internals\CS 1 Interviews\ 12 Manager](#) - § 1 reference coded [0.23% Coverage]

Reference 1 - 0.23% Coverage

When you are on call and beds are really bad the pressure to save money doesn't come into it at all.

Coded at CS 1: SWE; Workload

[<Internals\CS 1 Interviews\ 28 Doctor](#) - § 7 references coded [15.88% Coverage]

Reference 1 - 3.18% Coverage

So the medical patients instead of going directly to EMU come here. Now clearly we have constraints in terms of the availability of space, bed space and we have constraints in terms of the number of available nurses. Those beds space the calculations in terms of bed space and number of nurses are worked out on the basis that we are looking after the emergency department caseload and not the emergency department plus medical caseload. The medial caseload as I mentioned have a much higher acuity on average than the emergency lot and that puts huge pressure on the system in terms of bed space so we have to nurse patients in the corridor, we have to monitor patients in the corridor which is particularly vulnerable and we don't have the number of nurses to observe those monitors that we would have per patient should we not have the medical cases. So our pressures primarily relate to the acute medical take.

Reference 3 - 0.56% Coverage

The nurses however have overall responsibility for the entire burden of patients so when we have the acute medical take the nurses are under phenomenal pressure

Reference 6 - 3.40% Coverage

Well in terms of personal experience previously before we had the medical take if I was the duty consultant I had the time and the ability to supervise the care of all the sick patients. I didn't provide that care but I was able to be involved in the care and have my juniors report to me. Now when we have the medical take because I am doing effectively managing the department, trying to focus the resources in the most appropriate place, triaging patients, moving my doctors around, communicating with the management etc, I do very little of that which, and I am trained for the former, I am a bit trained for the latter but, so we've taken away the, what we lose in that situation is the consultant's opinion, the consulting with the consultant and I am sufficiently arrogant to believe that I bring some value added in terms of quality.

[<Internals\CS 1 Interviews\ 29 Doctor](#) - § 6 references coded [10.33% Coverage]

Reference 1 - 3.06% Coverage

I think that then has a knock on effect so the targets produce a need for additional to be done in the Trust which produces a capacity issue. So that work has to be done if it can't be done at 9-5 it will have to be done outside of working hours. So the Trust at the moment is doing an awful lot of additional work out of hours and at weekends.

Reference 2 - 2.47% Coverage

If you were to start where you were two years ago and say right we are going to get down to eight weeks nobody would do it because you just can't go from there down to there with as much, nobody would tolerate it but if you do it little by little by little you just find yourself running faster and faster and faster and faster gradually and you end up potentially starting to take on more than you should I think and what I don't know at what point people say hold on we've gone too far here we are pushing people too far. I mean you can look at I don't know staff attendance rates, sickness rates or you could look at the adverse events that we look at. I wonder if they are particularly sensitive markers my guess is they probably are.

Reference 4 - 2.25% Coverage

they can end up with you know literally space issues because they've only got so many bed spaces and people wanting to come out of theatre and there are physically no room for them to go into that's happened on occasions but more commonly than not it's puts a huge pressure on staffing levels particularly at the peak parts for recovery which tend to be later on in the day when all the lists empty out at five in the evening we are all going home thank you very much that's when recovery have big issues trying to get people onto the ward because all ten theatres are emptying into main theatre recovery at that time. So it, and it puts staff under a lot of pressure and particularly for example if you deliver a patient who needs airway support still has got a ...or a mask in or they need a jaw thrust or something they may not be in a surround to accept a patient because they are running round looking after other patients.

Reference 5 - 0.86% Coverage

I think in terms of general volume of work and things I don't think we, I can't think I mean recovery staff would put in an incident form in if they thought their staffing levels were dangerous. I've not seen any evidence of those through the governance committee either they are not putting forms in I don't know. I think they just kind of take it on.

Reference 6 - 1.10% Coverage

I think we have no data about that whatsoever. How do you determine when staff are overworked? I am not aware of any marker or ...chart on trust that says this is our level of staff pressures at the moment but I could say to you, you could look at the sickness rates and sayand things like that I think they are probably completely inaccurate so I don't think we have any measure of that.

[<Internals\CS 1 Interviews\ 22 Nurse](#) - § 9 references coded [11.49% Coverage]

Reference 1 - 0.27% Coverage

Money, not been mentioned, we spend what we've got to so staffing we are trying to get anything we can.

Reference 2 - 0.59% Coverage

We've been very pro active we realised that staffing is the centre of that so we have been pro actively recruiting to the point that we are hoping that we've actually got people ready to start posts before somebody even leaves.

Reference 3 - 1.48% Coverage

there's also the issue of what happens when somebody is sick on top of your normal staffing level or somebody is very confused and need so you know for each new thing that comes along the system is having to change and adapt to meet a changing need because a person falling over is at high risk of somebody who is very physically

unwell and previously I wouldn't say we had that emphasis you know before we had a confused patient walking around and you know they fall over and that's sort of at the point they fall over whereas now we are having to plan for caring for them.

Reference 4 - 2.08% Coverage

I think nurses at ward level become extremely uncomfortable so for instance last night we were very short of beds, we had an escalation plan, we had outliers identified, we had beds identified that were coming up that we could move patients into potential. The next level was that we would have to open a further area to care for patients of which there were no staff for so we had to risk assess in each division, medicine, surgery, orthopaedic about if we needed more staff was there any area or where was the safest area so the nurses would be safe that we could pull from and bring that plan back to the senior nurse and site management who were obviously going to be the ones here on site doing that but part of that was whatever plan I made I had to go and communicate that to the nurses on the floor

Reference 5 - 1.37% Coverage

So in this case we assessed that the cardiology pod had six trained nurses within that environment with most pods only having four trained nurses within their environment so therefore that was the pod that we would back fill with an unregistered nurse if we had to take a registered nurse out of there. So they weren't happy because you know they've got their own pressures they could see their own workload, but they appreciated that actually it had been looked at across the whole of the medical division about where was safest.

Reference 6 - 1.63% Coverage

For me it meant that I had to staff core areas, that I had to find out what staff were actually working in those areas and what their pressures were. Some decisions had already been made during the night of pulling staff out of areas to staff the extra areas that we've got open so by the time I came on this morning there were nurses in these extra areas moved from other areas someone had already made that decision.

So apart from CDU what other extra areas?

We've gotwhich is an area within the orthopaedic department which we've got 13 medical patients in being cared for by nurses supplied by the medical division.

Reference 7 - 0.88% Coverage

We have some flexibility in that the wards shut with the Noro virus have less patients in them so we've been able to move staff from those areas without compromising care or safety so we've had that buffer. We have also planned ahead and we are already pre booked bank nurses to come intoarea because we knew that was going to be open.

Reference 8 - 0.54% Coverage

It is an area that we are trying to plan for in the long term, we know that we are repeatedly opening these areas but in the short term it's, we spend what we've got to make sure we've got staff to provide care

[<Internals\CS 1 Interviews\ 04 Director](#) - § 9 references coded [9.24% Coverage]

Reference 1 - 0.60% Coverage

Now I think there's a question in terms of when we are as busy as we are you are by default overworking the staff and there's no option other than that. You know you can use that financial situation, you try and alleviate that but in the short term it's not possible.

Reference 2 - 0.10% Coverage

If the staff are not there they are not there.

Reference 8 - 0.86% Coverage

So those are there now the difficulty with you get a subjective what's going on and what's been quite interesting this year compared to last year is it feels, to say it feels calm might sound daft it doesn't feel calm in terms of what's going on but actually in terms of people coping and managing it and working through it feels calmer because actually we've got more staff in post.

[<Internals\CS 1 Interviews\ 16 Nurse](#) - § 1 reference coded [1.27% Coverage]

Reference 1 - 1.27% Coverage

We have within our division we have a staffing template which we know how many staff we should have on each shift and we know on a day to day basis if the staffing drops you know how we can, you know whether that's acceptable or not acceptable but we also have to think about what's actually happening on the ward at the time so we'll share staff around to make sure that every area is covered as you know as well as we can.

[<Internals\CS 1 Interviews\ 18 Nurse](#) - § 3 references coded [2.88% Coverage]

Reference 1 - 0.23% Coverage

over recruiting like we've done to give it a bit of a buffer, if we didn't have that it would have been really catastrophic.

Reference 2 - 0.15% Coverage

But we still haven't been able to fill every single person that's gone off sick.

Reference 3 - 2.50% Coverage

So nursing sickness, medical sickness, spread of further infection you know what I mean so how and cross infection in particular across the template. Staff welfare you know I just think when things like this happen and particularly you know its not happened for a long time actually that we lose that many wards in one go but just staff welfare.

[<Internals\CS 1 Interviews\ 06 Director](#) - § 1 reference coded [0.67% Coverage]

Reference 1 - 0.67% Coverage

so safety becomes more of an issue and staff I suppose we expect staff to overwork at this point in order to keep the patients as safe as possible

[<Internals\CS 1 Interviews\ 20 Nurse](#) - § 2 references coded [2.21% Coverage]

Reference 1 - 0.47% Coverage

I guess it's the impact that has on the staff when they are working with that pressure day in and day out I mean I present it as a very neat little.....oh well this happens, this happens, this happens it's all very typical but what I am not also saying is the impact that's having on the staff and it has a huge impact.

Reference 2 - 1.75% Coverage

EMU is a big burn out area a huge burn out. So for nurses to be able to sustain that level of activity day in and day out has a huge impact on their work./life balance and on their health really. So I am not an area that I can't, I cannot usually recruit into I have to do a lot of work in terms of recruitment it's not an area that's popular internally to move into.

Coded at CS 1: RP; DM; Capacity

[<Internals\CS 1 Interviews\ 28 Doctor](#) - § 1 reference coded [1.66% Coverage]

Reference 1 - 1.66% Coverage

The medial caseload as I mentioned have a much higher acuity on average than the emergency lot and that puts huge pressure on the system in terms of bed space so we have to nurse patients in the corridor, we have to monitor patients in the corridor which is particularly vulnerable and we don't have the number of nurses to observe those monitors that we would have per patient should we not have the medical cases. So our pressures primarily relate to the acute medical take.

[<Internals\CS 1 Interviews\ 07 Director](#) - § 2 references coded [2.11% Coverage]

Reference 1 - 0.97% Coverage

that's the safety valve if you like to make sure that the resources that we have available are sufficient to look after a group of patients even though it's a lot higher than normal we've released the pressure by not bringing in the ones that would normally come in and providing that the extra load is less than the elective patients coming in then it is possible to balance the management of the hospital in terms of the available beds and the resources but the sufferers are those that have a planned.

[<Internals\CS 1 Interviews\ 09 Manager](#) - § 2 references coded [1.77% Coverage]

Reference 1 - 0.93% Coverage

The CEO was in here one night at 12pm pushing beds round and I think that's ridiculous. I mean you don't pay the Chief Executives all that they get paid to push beds around at midnight and I know that looks good for PR for instance. Staff say – What's the Chief Executive doing pushing beds around. I mean I don't see that as a good use of the Chief Execs skills to be honest.

[<Internals\CS 1 Interviews\ 23 Nurse](#) - § 1 reference coded [0.12% Coverage]

Reference 1 - 0.12% Coverage

then we lost the day case unit because of bed pressures

[<Internals\CS 1 Interviews\ 08 Manager](#) - § 6 references coded [9.19% Coverage]

Reference 1 - 0.93% Coverage

if we took the attendances into the Emergency Department as a proxy for that I think when I started here you know 18 months ago if we had a few days where we had over 200 attendances we would be, you know we would be quite worried and feeling under pressure. Now we are consistently in the 230 to 260 and you know 230 now is a normal day, 260 feels busy so there's been you know a significant increase from that perspective.

Reference 2 - 1.96% Coverage

the first thing that we have done is very clearly not compromised the areas that are infected so those areas have got to be closed and they are quite clearly closed, there's no transfers out from them, there's no admissions into them so the patients can be discharged from them but we are not moving them anywhere else so they have been very much locked down and for us this week that's been kind of 8 ward's worth.

Reference 3 - 1.38% Coverage

We looked earlier in the week at all the elective operating to see whether there's any scope for doing that in different ways and certainly where we've had day case areas within specialties some of those areas have been used temporarily as in-patient areas to kind of accommodate patients pre-operatively and post-operatively until they can be put in the right places. We've been looking very much at you know any patients who are coming in for surgery today who are not having surgery until tomorrow putting those patients off and we've had a contingency plan about who is it that we would cancel if we needed to cancel patients.

Reference 4 - 0.78% Coverage

Some of our medical day case procedures have been happening in our, in Cherry Brook which is our kind of oncology area so blood transfusions and things like that have kind of gone in there. So we've looked at every kind of area of capacity within the organisation and been a bit creative about how we can keep work going but by doing it in different ways.

Reference 5 - 2.24% Coverage

We are still managing to put the work through but it's being done by some of the creative solutions that are kind of suggested around, you know and clearly there's a tension there but you know it's interesting because I think that when people know that the organisation is in the kind of situation that it's in at the moment they kind of pull together and it's interesting 'cos actually over the last two or three nights where we've had capacity and we've left, I've certainly left when I've left the building there's been a plan of opening additional capacity over and above the orthopaedic ward, when we've come in the next day none of that's been used and you know the discharges have come up but they've come up late in the day. The day cases have gone home and we may have utilised some of that capacity over night but I think people do kind of push things on at a slightly different pace once they, when it's clear that the organisation is under a significant amount of pressure than they might if that wasn't the case.

[<Internals\CS 1 Interviews\ 15 Nurse](#) - § 10 references coded [10.76% Coverage]

Reference 1 - 1.71% Coverage

So what we knew was that we did have beds in the system and I can't remember how many but it definitely equated to more than a ward but they were all over the place. So knowing that if we could cohort and create other beds could we then free capacity if we could get cleaners in knowing that on a Sunday that takes longer to do a terminal clean

Reference 2 - 0.75% Coverage

So what we looked at which we do it's a fairly routine thing, is there any other area in the hospital that you can create that place and "C" day ward in orthopaedics has piped oxygen and suction, we know they hate it but can more readily be made into an environment that's both safe and acceptable to patients, unlike some day case units where because they are trolleys

Reference 3 - 2.02% Coverage

the next issue was staffing if we were to do that how do you staff it and part of our problem was trying to keep what we call CDU open which is if you like an overflow but is not the best place to keep patients overnight so I'm not sure if you've seen it. It's a modular type temporary area. So what we came up with was actually a win/win would be if we could move those patients on CDU and compass those into the

numbers which I think brought us up to the 13 which would give us capacity at least for the scores on the doors i.e. the predicted numbers but the nature of “C” was that you could then expand it, it can go to 25 beds if for whatever reason your cunning plan you know all patients came through the door that we know happens and it is actually easier to cover safely with staffing that way than open lots of little areas with one or two staff and in order to open another five beds you might just need to support with a healthcare support worker rather than trying to find a whole

Reference 4 - 0.75% Coverage

We knew that we probably needed to open it to 13 to empty CDU and leave the capacity to get through ED and get us in overnight but if everything went wrong to move the capacity on that ward to open to 25. In order to do that though it sounds simple you’ve got to sort out your pharmacy store, staffing is the main thing and that took us probably about two hours to sort.

[<Internals\CS 1 Interviews\ 14 Manager](#) - § 3 references coded [9.06% Coverage]

Reference 1 - 4.51% Coverage

we’ve obviously got the two streams of work coming into the hospital of elective work and non-elective work and the elective work you can to an extent control but the extent to which you can control that is much reduced now compared to a position two to three years ago.

Just explain to me why that is. And the reason for that is the reduction that there has been in maximum waiting times so, and in that respect the introduction of base targets is driving how we manage the system so in other words now your maximum waiting times are about treating patients within 18 or 15 or 13 or 11 weeks of GP referral you don’t have a lot of time for patients to hang around on waiting lists and if you decide you are going to cancel elective surgery you are almost certainly compromising those waiting times and in addition to those waiting times you’ve also got waiting times for patient groups most notably cancer patients who have to be treated within 31 days of a treatment plan being agreed between the consultant and the patient. So your, your flexibility to switch elective work on and off is compromised.

Reference 2 - 1.93% Coverage

We’ve then got the non-elective work, the emergency work, and I guess your starting point for that would be how can you plan in for the emergency work because you don’t know what’s coming through the door. And I think that over the past few years in this hospital we’ve actually moved to a point where we believe that if that view is not the case you can make a reasonable plan for what’s coming through the door, so you, you can run models which predict activity based either on seasonal patterns but you can apply to those levels of admissions and discharges that have happened to get both a picture of what’s likely to happen through the year but also more immediately what’s likely to happen over the month ahead and very crucially what’s likely to happen on every day. So every day we go in, into the start of the day with a predicted number of emergency admissions which provides a real focus for the work that day.

[<Internals\CS 1 Interviews\ 19 Nurse](#) - § 5 references coded [21.30% Coverage]

Reference 1 - 0.33% Coverage

There are no days any more that we come in and there are beds.

Reference 2 - 1.70% Coverage

The story this morning was we have 13 patients who are out patients who come in for a mixture of procedures who will either stay overnight or will vacate the bed within a

few hours and I walked in this morning to bed 13 people into what would technically be three beds because that's how many I had definitely going home.

Reference 3 - 10.00% Coverage

So you have no beds to start with.

No beds to start the day with, so then I had to look at the ones who would have a procedure and vacate the bed, so I bedded those ones first so they had their procedure then we got patients ready whilst waiting for their beds and then some of the patients had to go into the bed that the first people who were diagnostic and went out of the bed had vacated, do you understand that?

So what the patient who is having their procedure they vacated a bed which you would then use again.

That I would use again for somebody else who was having a procedure that would then stay in the bed overnight. So this morning any beds that I got were then blocked because people who stay overnight once they are in the bed the bed's gone because you know they've got the bed and it's booked. So then to start the afternoon I had no beds at all again, fortunately some people are being discharged but it means people are waiting you know in waiting rooms waiting for beds that you are juggling to see whether they are going to come up or not. We also have to, because I govern the beds for the whole of the cardiology template you then have to think all the time with, you know with because we do a primary angioplasty service as well so somebody comes in and they go directly through to the lab, if coronary care is full you then have to think where on the template can I bed somebody to make a bed for the emergency if there is one and you always have to have a plan in the back of your mind if an emergency comes through the door this is what I will do and then you've got [Ward] "A" beds as well but unfortunately today there aren't any you know early in the day coming up if any at all. So it's just working out where you are going to bed people and if you can bed people.

So that was today's example, how often does that happen?

Every day, every day.

[Internals\CS 1 Interviews\ 13 Manager](#) - § 1 reference coded [0.10% Coverage]

Reference 1 - 0.10% Coverage

we just don't have the physical capacity to be able to do the work that is there

[Internals\CS 1 Interviews\ 20 Nurse](#) - § 9 references coded [13.85% Coverage]

Reference 1 - 1.92% Coverage

...overwhelmingly the momentum is driven by demand and capacity and flow and you know you learn that concept very quickly when you work on EMU. We have to manage the admission process through the EMU we take those from the GP and also the emergency department. At this time of year we are moving into the 55/60 every 24 hours, that's a large number coming through, we are a 31 bedded unit. So for the first part of the day when we are taking calls we are actually having to manage the emergency take ourselves because the wards won't be discharging at that time of day and they won't necessarily have the beds so therefore our, the EMU acute physicians have to really get busy with discharges. Now they have the first whip if you like in terms of they've probably got the best client group to discharge because a lot of them have come in overnight or night before, evening before and so we can aim to expedite some discharges that way.

Reference 5 - 3.14% Coverage

Okay I can tell you actually because I went in and I was given that information so I've got 31 patients on EMU no beds available and two patients to come in. We've also got 13 beds out on the orthopaedic ward which is pre-assessment and we have taken over 13 beds out there which is also EMU and we've also acquired but we are not managing 6 beds in the CDU, clinical decision making unit, which is I don't know if you have been down there but it's a little area in a portacabin which is tucked away in the emergency department so that's an awful lot of beds that EMU have suddenly acquired geographically

Reference 7 - 1.33% Coverage

Well at that point its clearly escalated and the site practitioners are talking you know to start ...and so on and so forth and the divisional managers are in and the discharge team are in and they have to start making very challenging decisions I guess around elective work and you know freeing up day case beds and you know and elective beds and stuff like that and that's when the other six beds in the emergency department so last night...so yes they are always looking for capacity where can we you know leave space for clothes if it's a day case area so now we will have to open it as an in-patient area which then has a knock-on effect for the elective work for the next day. So there's a lot of decisions made with the on call teams after 4 o'clock in the evening and a lot of the decisions have to be made at a very senior level before we can act on anything.

Reference 8 - 0.80% Coverage

you've got these elderly frail patients because that's really what the A&E medicine is about you know sat on trolleys and you've got four hours with which to bring them into an area. Now they do have beds and they do transfer their patients on to beds so that they are comfortable but it's not the right environment for them in that, you know and they are confused and unwell to be in the emergency department and ED is ED and it has to function in terms of its minor injuries, majors, resus and we are clogging it up with emergency care.

Reference 9 - 0.33% Coverage

it wasn't an operational area it was pre assessment area so even now we are still you know we haven't got the right stores in place and the pharmacy's not quite right because of course it's not an established in-patient area.

[<Internals\CS 1 Interviews\ 17 Nurse](#) - § 6 references coded [6.23% Coverage]

Reference 5 - 1.64% Coverage

the ones that have to be fast tracked to surgery obviously they are the ones that you know you have the list this big and you have that many beds and the patients are coming through for their fifteen / eighteen weeks the pressure is on you know and you know these patients need to come through the system but you just wonder if you will ever come to a state when you saturate when you actually reach the point where there is nowhere else to go. I can't imagine it happening because once we've opened up all the day case areas I can see us going into endoscopy and you know we will be using every physical space

Reference 6 - 0.51% Coverage

You know you've got your infection control looking at every option of cohorting patients that do have the virus to try and free up any little bit of capacity that's safe to do so that you might look at cohorting two wards of patients providing you've got enough beds to do that.

[<Internals\CS 1 Interviews\ 24 Doctor](#) - § 7 references coded [7.35% Coverage]

Reference 3 - 1.18% Coverage

It also you know reputation wise it's bad for the hospital it just causes so much upset cancelling operations so where we can we try and not do that and similarly we try and not cancel people who are due to come in for coronary arterial grams or anything that involves family disruption and anxiety but the price you pay then is people sitting about in corridors all day you know which is undignified and uncomfortable and perhaps not getting done at the end of the day after all anyway but I think it is, it is right to try and get them in if you can but it does produce a whole lot of stress and produces an unsatisfactory working environment

Reference 7 - 1.47% Coverage

I do think there is a genuine capacity problem here you know I don't think that would be particularly contentious but I can understand the anxieties that whenever you open beds they get filled but you know I can't quite see you know you can become more and more and more efficient but there has to come a point where demography just will overwhelm the place because it's you know it's sort of experiential rise of the very group who generate the longer stays the greatest morbidity and I don't see how we could actually go on much longer without some sort of more definitive increase in capacity. Now obviously the community hospital theatres might offer us a great opportunity as long as the surgeons are disciplined enough to actually just not admit any day case stuff here get it all done out there.

[<Memos\Memo 041109 CEO hierarchy of decision making during bed crisis>](#) - § 1 reference coded [2.03% Coverage]

Reference 1 - 2.03% Coverage

So you've got to cancel and free up the beds that, so we've got that, at the moment we've got three surgical wards orthopaedic and surgical wards transferred across the medical wards.

Case Study 2

Coded at CS 2: SWE; Patient Safety

[<Internals\CS 2 Interviews\ 01 Director](#) - § 4 references coded [8.45% Coverage]

Reference 1 - 1.55% Coverage

all of the patient safety stuff obviously and MRSA and C-Dif and all of the infection control targets they are managed and in fact they are listed first, the patient safety stuff, on our performance dashboard the patients' safety issues come first but people are clear that they manage all things and that they have to juggle all things and I guess that's what the complexity of the NHS is isn't it [laughter] you know it just comes with the territory.

Reference 2 - 3.18% Coverage

what other things that are on there that you count as awareness of patient safety?

Well we are briefed, because our numbers now of C-Dif and MRSA are so small, they are every single one is reported through a route cause analysis. So we get a good deal of information on all of that. Noro virus outbreaks are briefed to every head of department as they happen so that we have a standard email alert system that goes out for that. Incident reporting is the methodology that we would use for any security or aggression incidents or any sort of patient safety incident

[<Internals\ CS 2 Interviews\ 03 Manager](#) - § 3 references coded [14.49% Coverage]

Reference 1 - 9.46% Coverage

We have just in the last six months started to produce a first cut of a patient quality and, patient safety and quality indicators' report. This is the first time we've brought together such a wide range of information within the Trust. We prepare that report on a monthly basis it's shared with our finance and performance committee internally which is the executive and non executive directors. We share that with the primary care trust through the commissioning and contracting route and it's a means of us pulling together certain key indicators and being able to then demonstrate the trend performance over a period of time against those indicators.

Yes so what type of things are you measuring?

It picks up quite a wide range of indicators at the high level. We are picking up hospital standardised mortality ratios, re-admission rates, we pick up information on incidents, serious incidents, general incidents and information on the levels of complaints that are being presented the organisation from patients and members of the public. Information on things like pressure sores, we've just started now to collect data on VTE, information on accommodation so mixed sex accommodation so where we breach standards of accommodation that we pick up so we incidence form and analyse that, determine whether that is an acceptable level in the context of the numbers of patients coming through the system or not.

Reference 2 - 1.84% Coverage

They would all be reviewed monthly. There are some instances for example on the mixed sex accommodation we measure that on a weekly basis, we get weekly data through and we provide an interim report to the executive team each week so that they can see whether we are on track with that, but most of them are reviewed monthly.

Reference 3 - 3.19% Coverage

I see the number of medical outliers on a daily basis I don't specifically measure and report but basically the medical outliers are included on a daily bed state email that comes out every day. We don't provide additional information on medical outliers. ...I am not aware that anything is actually presented to the executive team or to the Board on medical outliers.

[<Internals\ CS 2 Interviews\ 05 Director](#) - § 11 references coded [32.14% Coverage]

Reference 2 - 5.61% Coverage

So then I was asking you about the information you get at Board level in terms of patient safety. What things do you get? Well we get a very good report on infection control from the Director of Nursing. ...At Clinical Governance we have SUIs reports and there's never been one in the last three or four years on patient safety issues all the SUIs have been something else, mostly clinical stuff you know.

Reference 4 - 1.26% Coverage

I am just wanting to clarify because when you said clinical issues as SUIs what

type of clinical issues are those then? Oh these are things like deaths in maternity things like that that you couldn't really say was a, there was a failure of the organisation on safety grounds.

Reference 5 - 4.52% Coverage

have there been failures in other grounds then do you think?

Well I mean the SUIs are failures aren't they so I mean there have been, oddly enough there have been three deaths on cataract operations in a very short space of time but again I mean the investigation on that was just coincidental you know it wasn't a trend at the end of the day and the other thing is I get, you see these incident reports and we are encouraging the staff to fill in incident reports and we think if we are going up on that that's a good thing because there's a, you know people are reporting things and we only take the incident reports higher up to the Board or Clinical Governance if a trend is emerging. You know if it's just a one-off then that's it and again patient safety has not been picked out as a trend, patient danger really [laughter] yeah has not figured as a major trend in the organisation. So I mean those are the sorts of things that assure me at Board level that there's a system underneath.

[<Internals\ CS 2 Interviews\ 07 Director>](#) - § 5 references coded [13.06% Coverage]

Reference 1 - 7.00% Coverage

what other kind of patient safety issues which are of concern and are measured and monitored? Well I should have brought a list. [Laughs] well I suppose you can go from the very, very specific things which is like clinical practice things such as have they got a line in, when was it put it, is it dated, is it timed, is it checked, does it come out, so there's very practical things around lines, catheters, insertions and things like that. There is a kind of so process driven they are things like obviously hand hygiene you know which impacts on everyone, there's your equipment have you got the right equipment, the right skills and competencies of the nurses caring for the patients, have you got the right moving and handling staff so, not that that's about health and safety as well but if the patient moved appropriately there's very specifics we do which is around obviously more infection control stuff, lots of measures in ITU looking at ventilator requirements, pneumonias, VTE risk assessments, compliance withcompletion I mean I could, do you want me to carry on?

...[Laughs] I've got to go through my list now in my head and remember them all. We do early warning scores, audits, we audit all the cardiac arrest calls, and with the other much more strategic stuff is we've, every clinical incident I review. I get all the clinical incident forms and then we would do kind of an overview in terms of themes every three months and pick up any more themes the massive stuff that comes out and may be follow that up with a task and finish group if there is any specific things around that. I mean patient transport we look at the safety of that.

Reference 2 - 1.38% Coverage

Well the level of harm I mean I can tell you what it was for last year through incident forms and I know that that's a small percentage of the totality of reporting, 603 patients from our review last year were harmed but that's, they reckon it's what only 3% of you know what you might get so you can extrapolate up to thousands.

Reference 3 - 0.75% Coverage

Now rightly or wrongly what they will focus on is the MRSA infection which is probably you know it's only two patients this year for us but we've probably got 102

that have had a urinary tract infection. So I don't think they've got the degree of harm.

[<Internals\ CS 2 Interviews\ 08 Manager](#) - § 1 reference coded [7.09% Coverage]

Reference 1 - 7.09% Coverage

What's sort of concern and what's measured in terms of knowing what's happening to the safety of the patients? I mean I think the concern is that when the organisation is under pressure everybody is like on a hamster wheel. You know the acute physicians are seeing turnover of patients going through MAU and it's about how we protect their ward round and their time to make sure that patients are being properly managed. ...So I think there is something about what pressure we are putting the consultants under. The pressure we are putting the juniors under when they've got fifteen patients on their own ward and then they've got another twenty-five outliers and how those juniors get round and review. I think the pressure on the consultants between elective work and in-patient work when we are saying you know you've got to get your waiting time to four weeks you know you've got to make sure your clinics are full and then the balance between and by the way you've got to make sure that you know all your patients are being discharged and they've got plans.

[<Internals\ CS 2 Interviews\ 09 Doctor](#) - § 2 references coded [9.00% Coverage]

Reference 1 - 2.06% Coverage

Well for the patients they say they wait longer, they wait longer for everything. They wait longer for actually get to a cubicle sometimes, they will wait longer for their initial assessment, they will wait longer for pain relief, they will wait longer for investigations and they will wait for is they are going to be admitted a member of the in-patient team to come and see them and make that decision and during that time there is a potential for their condition to deteriorate. So this is not an issue of convenience or you know high quality service and you know people are seen promptly which is what people would like you know it's an issue that people can deteriorate while they wait.

[<Internals\CS 2 Interviews\ 10 Nurse>](#) - § 2 references coded [2.68% Coverage]

Reference 2 - 1.90% Coverage

I mean clearly sort of patient falls is one of the main risks sort of within the area because we have quite high levels of confused, agitated patients that are elderly that have, they are taken out of their normal environment into our environment which is very busy and obviously they start to get agitated and the risk of falls is greater.

[<Internals\CS 2 Interviews\ 11 Nurse](#) - § 6 references coded [12.67% Coverage]

Reference 1 - 1.73% Coverage

I think in my mind I am always very clear that my message absolutely is patient safety. There are times when, with patient safety, I think we do need extra staffing or we do need to put the money where the need is. There are times when I think we can work differently and work smarter to achieve that same patient safety. When it's working smarter I am very comfortable that we need to get on and do that, when it's actually that we need the money because there's no other way to manage this situation I do feel some conflict because sometimes it's not easy to actually do that.

Reference 5 - 0.47% Coverage

The medical cover is an issue for us. A real issue because we have to be safe but equally you have to come in on budget so that is a huge challenge for us.

[<Internals\CS 2 Interviews\ 14 Nurse](#) - § 2 references coded [3.68% Coverage]

Reference 1 - 1.94% Coverage

Because I think patient safety you know staffing numbers relates to the safety of the patients really that I think you need the patients there, the staff there, to ensure that patient safety is happening really because without the staff you can't monitor that and you can't, if you've got patients climbing out of bed and falling and you've not got enough staff then obviously that's going to potentially happen and those patients are then at risk of you know injuring themselves really.

Reference 2 - 1.73% Coverage

other than falls what other patient safety issues are you, do you worry about? The infection control is a big thing if staff are rushing between patient to patient because they are short staffed they are not going to be doing their hand washing and therefore spreading infection especially you know with what we've got at the moment it's you know hand washing is every important but if they are rushing or if a patient is climbing out of bed and they've got to run to get them and they've just dealt with another patient and it's you know the risk yeah the spread of infection is very high there.

[<Internals\CS 2 Interviews\ 15 Doctor](#) - § 3 references coded [13.08% Coverage]

Reference 2 - 3.94% Coverage

have you experienced or seen anything going wrong with patients as a result of them being on an outlying ward? Well I mean I think it's all, I think it's historical that we always think that patients aren't being managed as well on a surgical ward. We have a lady at the moment with breathing problems who probably would have benefited from being on the respiratory ward but because its shut she can't be there and although she is not deteriorating she's not really getting any better and it might be as a result of the fact that they don't give the medication quite at the right time that the nurses would know on a respiratory ward. I think because we are quite careful about going to see and most of the teams are but going to see every one of the outliers that you do pick up what's going on.

Coded at CS 2: SWE; Targets

[<Internals\CS 2 Interviews\ 01 Director](#) - § 2 references coded [8.46% Coverage]

Reference 1 - 4.31% Coverage

How do you know where you are on targets and how often do you monitor that? I am fairly obsessive about that [laughter] I get daily reports on some of the stuff. So I get real time information on, four hour target, I get real time information of any breaches in particular, I get weekly information on where we are with our RTT position and all of the bookings team manage that on a real time basis and there are weekly very operational meetings which I don't attend unless something is going badly wrong around RTT and around cancer waiting times. So the MDT coordinators and people at that level so the senior booking clerks will be part of that meeting to work with the patient level data. We oversee it on a, I have a monthly what we call a planning contracting and performance management meeting which is, which I chair and that's all of the general managers are in attendance with the performance management teams, IT and finance to try and bring all of the performance measures that are contracting issues together so we review that on a monthly basis and there's a weekly PTL that goes to execs which includes ambulance handover breaches for our targets and the progress of the week for RTT. So there's quite an in-depth scrutiny

about that too.

[<Internals\CS 2 Interviews\ 02 Director](#) - § 1 reference coded [0.11% Coverage]

Reference 1 - 0.11% Coverage

we always deliver every target

[<Internals\CS 2 Interviews\ 03 Manager](#) - § 5 references coded [14.98% Coverage]

Reference 1 - 3.06% Coverage

We have in place a performance management framework within the Trust whereby data is analysed at the lowest level and information from that data analysis is then communicated through the Trust at appropriate levels right the way through to the Trust Board on a regular basis. Different levels of information at different or different detailed information at different levels within the Trust. So the Trust Board will get a high level view. With further detail and exceptions, whereas the directorates at the lower level and various management teams may get a much more detailed view.

Reference 2 - 4.96% Coverage

And what's the kind of timing of these reports, is it on a very regular basis? The minimum [pause] would be on a monthly basis where everything is performing as it would be expected but there are a number of areas that we review on a weekly basis, there are some areas that we review on a daily basis.

...Yes information is assembled daily for example on the A&E four hour wait performance and that includes information on emergency admissions and ambulance handovers and that's disseminated daily. Information is disseminated weekly on the level of elective referrals coming into the Trust on the outpatient activity, elective activity, outpatient waiting lists, elective waiting lists all versus the planned positions.

Reference 3 - 3.95% Coverage

We may only disseminate it monthly under normal circumstances but by reviewing that information on a weekly basis I can raise exceptions up the line if I need to. So if I see anything that's going astray that is in let's say one of our more routine areas where we rarely have problems I can still get that escalated and attention given at the appropriate levels within the Trust.

So do you have an example of having had to do that recently? Yes on cancer waiting times we have been struggling to achieve one of the cancer waiting time standards so although we generally only report that as a monthly position where we are normally on track where we've been off track then we've been reporting that through to the directorate teams on a weekly basis and having discussions with various members of the executive team. Just to make sure that our plans are strong and appropriate and delivering the improvement.

Reference 5 - 2.04% Coverage

Cancelled operations information is measured generally on a monthly basis but for example with the level of cancelled operations we've had going through in the last few weeks we've been measuring that several times a week. So it's all recorded on our cancelled operations database whereas I would normally seek that information on a monthly basis I've been pulling that information off every few days just to check what numbers of cancelled operations are going through.

[<Internals\CS 2 Interviews\ 04 Doctor](#) - § 1 reference coded [1.49% Coverage]

Reference 1 - 1.49% Coverage

No I am more likely to find out that they are not meeting their targets in situations where the hospital is full and there's a crisis if you like. So I am more likely to become aware that those targets aren't being met in crises.

Not at other times when things seem to be okay.

[<Internals\CS 2 Interviews\ 05 Director](#) - § 1 reference coded [4.05% Coverage]

Reference 1 - 4.05% Coverage

I am very happy and comfortable with the information we receive on performance.

[<Internals\CS 2 Interviews\ 08 Manager](#) - § 1 reference coded [4.69% Coverage]

Reference 1 - 4.69% Coverage

How did you do on the A&E department target during all this busyness? We had very little flexibility, we'd had quite a difficult time sort of around the autumn where we used up a lot of our lives, we put in a fairly major action plan around how we were going to bring our targets back and through the autumn things seemed to get better but because of the numbers of breaches on a few days because here because our numbers or emergency admissions although relative to us are high we haven't got much room or much flexibility and we have kept our head about water until the last week or so and we are really struggling again. At the moment we are below 98% so we are working.

How far below? When I looked just now it was 97.2% or something but still that does not give us a lot to year end. I've met with the A&E consultants, I've met with the executive team because they are at the moment conflicting well not conflicting but there are financial concerns and A&E target concerns and we have discussed with the execs today neither of which we can afford to miss out on.

[<Internals\CS 2 Interviews\ 10 Nurse](#) - § 2 references coded [8.70% Coverage]

Reference 2 - 3.89% Coverage

I mean generally A&E will constantly inform us and say this patient is coming up to their breach time sort of can we move them up and we try to work as effectively as possible sort of with them so either create the bed or if the patient is well enough we utilise our waiting area because we've got a clinic running with the waiting, next to the waiting area. So there is a senior registered nurse in there so there's the potential to continue to observe patients there and then bring them into the bed when we are able. Or review them in the clinic that's the other opportunity.

[<Internals\CS 2 Interviews\ 11 Nurse](#) - § 2 references coded [6.31% Coverage]

Reference 2 - 5.11% Coverage

The position is we want to go forward as a Foundation Trust we want to continue the good work that this organisation has done. So the message was very clear you've got to meet your targets.

[<Internals\CS 2 Interviews\ 12 Doctor](#) - § 2 references coded [2.87% Coverage]

Reference 1 - 2.02% Coverage

there is a lot of pressure and we have a few sort of crisis points here. Cardiology and gastroenterology are two specialities closest to the targets and closest to not hitting the targets so we do put on extra clinics and extra endoscopy lists sort of in addition to the standard work programme to facilitate us hitting the targets.

Coded at CS 2: SWE; Finance

[<Internals\CS 2 Interviews\ 01 Director](#) - § 1 reference coded [5.52% Coverage]

Reference 1 - 5.52% Coverage

On a more operational level we run our monthly performance reviews which is our opportunity to sit down with the divisional teams and look at their monthly I and E position and they cumulative year to date and their year-end forecast position and we will have a number of not just the budget information but we would have a number of indicators that we would look at around agency staffing usage, vacancy control information and no vacancies get approved for filling without an exec team, they all come to the exec team every week and similarly agency usage is only approved by myself of the Director of Nursing and out of hours on call etc. We have, we regularly review short and long term sickness levels and those sorts of things as indicators to underpin what's going on with the pay position, the variable pay position. Because our experience over the last couple of years has been that that is the thing that affects our I and E the long and the short of it is so if you are not in control of your variable pay then you are not in control of your budget.

[<Internals\CS 2 Interviews\ 02 Director](#) - § 1 reference coded [0.36% Coverage]

Reference 1 - 0.36% Coverage

So one is the most important? I think it is still 'cos it's the legacy of the financial deficit.

[<Internals\CS 2 Interviews\ 05 Director](#) - § 3 references coded [5.80% Coverage]

Reference 2 - 2.37% Coverage

In terms of the reports that the Board get to back up the state of the finances.

Yes they are detailed and they are, we have a very lengthy finance committee meeting and go through the whole thing in great detail and at the Board it's virtually a rubber stamp situation because we've gone through it all before.

Right. Who is on the Finance Committee is it? All the, there's all the non execs so I mean. **So it's a mini Board meeting.** It is, it is yes. **So that focuses specifically on the financial position.** Yes.

Reference 3 - 1.54% Coverage

what are the kind of key priorities for you as a Board? Well finance is obviously we keep that under control.

[<Internals\CS 2 Interviews\ 08 Manager](#) - § 1 reference coded [4.09% Coverage]

Reference 1 - 4.09% Coverage

There are issues around lack of income, if you do cancel surgery there's the loss of income, there is the pressure on targets to achieve waiting times so then one of the consequences of that is Saturday operating to catch up on the cancels which costs. The use of agency when we have to put up additional beds, in medicine we've got pressures on our middle medical rota at SPR level so we've got a real problem with covering every night with a senior doctor so we are having to do quite a lot of work around getting that covered and that can sometimes mean, because I've got no alternative, the use of expensive locums. So that busy period has quite a significant knock-on effect

[<Internals\CS 2 Interviews\ 11 Nurse](#) - § 4 references coded [12.96% Coverage]

Reference 1 - 2.77% Coverage

...when it's actually that we need the money because there's no other way to manage this situation I do feel some conflict because sometimes it's not easy to actually do that. Having said that if I absolutely believe that that's the right thing to do then I will do it and I am very happy to stand up and be counted for that as I do so every month at the finance and performance. But I do feel we have a duty to come in on budget and work within our financial constraints where practically possible to do so.

Reference 3 - 5.11% Coverage

The position is we want to go forward as a Foundation Trust we want to continue the good work that this organisation has done. So the message was very clear ...we've got to reduce our overspend. So that message was very clear and we will do our level best to do that primarily in medicine we have done extremely well this year performance wise and finance wise. The wards were notoriously or is the biggest overspend I am very proud to say that I have brought all of my wards in on budget other than the completely accountable expense where I've booked extra nurses and resources for specialist patients and I would stand by that because that is my patient safety focus and that's legitimate. Where our spends have been bad in medicine this year and I can't speak for the other directorates because they are all different, would be medical cover obviously nationally there are shortages in middle grades and consultants and we feel it here in our locality so we've had a lot of locum use and that has literally our overspend is medical cover and on non-achievement of our CRES because our CRES that we put forward was to reconfigure the wards and we weren't allowed to do that because we had to achieve the same set agenda and there just wasn't the capacity to do both. So I think if you were to look at our performance overall I think we've done extremely well. The medical cover is an issue for us. A real issue because we have to be safe but equally you have to come in on budget so that is a huge challenge for us.

Coded at CS 2: SWE; Workload

[<Internals\CS 2 Interviews\ 01 Director](#) - § 3 references coded [8.05% Coverage]

Reference 1 - 1.77% Coverage

we would have a number of indicators that we would look at around agency staffing usage, vacancy control information and no vacancies get approved for filling without an exec team, they all come to the exec team every week and similarly agency usage is only approved by myself of the Director of Nursing and out of hours on call etc. We have, we regularly review short and long term sickness levels and those sorts of things as indicators to underpin what's going on with the pay position, the variable pay position.

Reference 2 - 3.63% Coverage

Yes, yes because we have now quite good information through our nurse rostering systems and through ESR so we can look at down to individual staffing levels about overtime.

So just explain what ESR is. Electronic Staff Rostering System and our electronic payroll records so we can bring the two things together and look at how efficient the rosters are, where there have been gaps. So if we haven't had an appropriate person in charge at Band 6 or Band 7 it will highlight that and we can, we can force it to make judgements. So we set appropriate boundaries, for example we would say that we would always expect a Band 6 or a Band 7 to be on, on certain shifts and if you go

outside of that through the E-Rostering then what actually happens as well as what's rostered then it flags it. So we have quite a good range of indicators that show what sort of pressures the clinical staff are working under at that level.

Reference 3 - 2.66% Coverage

So tell me what would happen then if there wasn't a Band 6 when there should be a Band 6 what type of actions get taken? Well on a day to day basis that would obviously be the responsibility of the Ward Manager and the Senior Nurse but if that's consistently happening on a ward that would be flagged up to myself and the Director of Nursing well in fact the report goes to all the Exec Directors but it would be "C" and I who would take action about that because it was, it would obviously either mean that there was some sort of performance issue with the Ward Manager not running her rosters properly or some sort of sickness issue that wasn't being picked up which meant that people were going off at a senior level and obviously that would indicate stress isn't it.

[<Internals\CS 2 Interviews\ 03 Manager](#) - § 1 reference coded [4.44% Coverage]

Reference 1 - 4.44% Coverage

Within my team we do prepare workforce information which has some financial detail associated with it and that financial detail comes from the finance team as well.

Yes so tell me about what information you have in workforce terms? We produce a workforce monitoring report that picks up information in respect of, at an overall level how far the Trust is from its budgeted position in terms of overspend or under spend. We pick up information on the amount of overtime being used within the Trust, information on the amount of agency staffing or bank staffing that we use. We just track these things to make sure that we can understand if trends are changing that we understand the reasons for those changes and if there are adverse changes then we bring that to the attention of the right people at the right time. We pick up information on staff sickness rates and monitor that and produce more detailed information at individual department or directorate level and that goes out across the Trust on a monthly basis.

[<Internals\CS 2 Interviews\ 04 Doctor](#) - § 1 reference coded [1.15% Coverage]

Reference 1 - 1.15% Coverage

I think we probably are adequately staffed on the medical sector to deal with the workload but because the burden or the sometimes it feels the ...responsibility is with me as a consultant on the shop floor as it were and you can be, the urgency with which we need to crack on and move things faster falls to me as well

[<Internals\CS 2 Interviews\ 05 Director](#) - § 1 reference coded [4.07% Coverage]

Reference 1 - 4.07% Coverage

what about staffing do you know how you fare on the staffing side of life? That's a bit more mystical I think. I mean we look at the graphs I mean all I can say is that for some reason every time there is a rise in expenditure on staffing, which must mean there are more staff somewhere. I know there are pressures but again we've asked for a lot of information graphs, new graphs and things and I mean the Director of Human Resources two years ago had to leave the organisation there was a weakness there and we were getting duff information basically, or the previous Board was but no I, we are getting the information and probably, and we've started a, there's a workplace planning organisation, there's a committee now looking at workforce planning for the future. And I think it's becoming more on the rails but it's certainly not as efficient

information as we get on the financial side.

[<Internals\CS 2 Interviews\ 07 Nurse](#) - § 3 references coded [4.79% Coverage]

Reference 1 - 2.92% Coverage

...strategically I know the levels of staffing for each ward so they have an established level of staffing for each shift depending on what shifts they do in a day. We have a cross wide risk assessment tool that's used in all the wards which is called a Workload Assessment Tool which at the beginning of the shift a, the ward would rate their patients in terms of numbers they've got in total in their beds assuming they are not full, the care needs that they have so how many people on IV's, how amber patients i.e. the sicker patients which would allow on a day to day basis to get an objective rather than a subjective sense of where the highest pressure points are which we use as a decision making tool in terms of if we have gaps in our staffing where we then re-deploy to rather than he who shouts loudest gets is trying to be but it's never going to be black and white but it tries to give some objective measure. So strategically I know what numbers there are on a day to day basis.

Reference 2 - 1.54% Coverage

We also have an electronic roster system so that at any point in the day I could go on and look at any ward to see how staffed they were and the other way I probably get a better sense of it is to spend time on the wards and I do that and we do a patient safety walk around here which I know lots of other places do and quite frequently the questions will come up about staffing and you get a sense from that but out of that that picks up my point earlier is that I know professionally that we are well staffed in our areas

Reference 3 - 0.33% Coverage

So if we have shortages on the ward I do generally know about it because all the agency request go through me.

[<Internals\CS 2 Interviews\ 10 Nurse](#) - § 1 reference coded [2.61% Coverage]

Reference 1 - 2.61% Coverage

So I feel that the unit has a very flexible workforce which is absolutely paramount for the type of work that we do and it's not a case of when it gets busy it's busy all the time so that the quiet periods are quite rare you know you generally expect to be very busy and then at times it can be extremely over busy and that is generally reasonably manageable as long as we have the staffing in place. Once sort of we haven't got the staffing because I think personally I feel that we're staffed at the minimum levels that we can manage the workload safely. Once we drop below that I feel that things potentially can be compromised and we obviously then have to re-prioritise what we do.

[<Internals\CS 2 Interviews\ 11 Nurse](#) - § 2 references coded [2.38% Coverage]

Reference 1 - 1.54% Coverage

I think we have a very good understanding here, we've done a huge amount of work on establishments. It is actually a personal interest of mine and I have done a lot of corporate work on establishments workforce analysis and financial involvement of that. So I think we've got it in a good place but I think the problem we still have within the NHS is our sickness levels are high and if you can't cover sickness and short term absence then you then have got wards running short and that's a long term problem within.

[<Internals\CS 2 Interviews\14 Nurse](#) - § 2 references coded [5.41% Coverage]

Reference 1 - 1.94% Coverage

Because I think patient safety you know staffing numbers relates to the safety of the patients really that I think you need the patients there, the staff there, to ensure that patient safety is happening really because without the staff you can't monitor that and you can't, if you've got patients climbing out of bed and falling and you've not got enough staff then obviously that's going to potentially happen and those patients are then at risk of you know injuring themselves really.

Reference 2 - 3.47% Coverage

We have a working workload assessment tool which we use every day and that gets completed by the night staff and they put on there all the staffing levels and then we have the acuity of the patients, how many patients we've got on IVs, how many confused patients, dependent patients. So you work out the scoring and it gives a green, amber or red and the staffing is the same you have either a green, amber or a red. So if you have green staffing and amber patients potentially you can cope with that but if you've got amber staffing and amber patients you might struggle so therefore you can highlight that you know you've got both amber and you can say what you've done so you can look at your staffing levels then and say right okay I've got five people on you know I could perhaps move one, move person, we need people in the morning because we've got 25 dependent patients and therefore we will move somebody from the late shift to the early to do the washings rather than the afternoon. So we can look at that daily and that goes to the clinical site managers and obviously they are managed by "BC" as well so she looks at that as well but they are used across the whole of the Trust.

[<Internals\CS 2 Interviews\ 15 Doctor](#) - § 1 reference coded [4.36% Coverage]

Reference 1 - 4.36% Coverage

Well I think we just, we need more doctors and I think the nursing staff would probably say ideally they would want more nurses we just need more people actually on the wards and we don't necessarily needs lots of people in senior reg positions we need actual sort of SHOs, house officers to be doing the sort of day to day duty with people supervising and I think at the moment we are really pushed here we are all doing a lot of sort of extra shifts and that's taking us away from our day jobs and it's leaving a lot of our junior people to look after patients without the sort of I think background knowledge and skills. And I don't think that's really about that's not really about money it's just about getting people but for this hospital we are just not attracting enough doctors.

Coded at CS 2: RP; DM; Capacity

[<Internals\CS 2 Interviews\ 07 Nurse](#) - § 1 reference coded [2.07% Coverage]

Reference 1 - 2.07% Coverage

last week we had some extra beds up so what I had to do was to obviously look at using the pool differently or re-deploying people.

So where did you put the extra beds? We put them up last week in "Ward P" which is female ward and it has day case work so in the end we put beds up in there so they are in a ward. **So there's a day case facility on a ward became an inpatient beds.**

Yes, yes. **How many beds was that extra you put up?** Last week I think it was up to 24, it would be about. **And is that a fairly common occurrence for you to have to do that?** Not that common. It was more common last winter. **Right.** But we've opened a winter pressures ward this year which I think has led it to be less.

[<Internals\CS 2 Interviews\ 09 Doctor](#) - § 3 references coded [3.12% Coverage]

Reference 1 - 1.36% Coverage

So they are all brought into the corridor waiting for a space and the impact for them is, I mean I think that is an appalling situation because from a clinical point of view there is not good monitoring in the corridor obviously but more importantly and from a privacy and dignity point of view you know people who are potentially distressed, vomiting, in pain, in a corridor with people walking past them there is absolutely no sense of privacy at all.

Reference 2 - 0.89% Coverage

When things are really bad we can certainly have five or six patients queuing and hopefully they wouldn't queue for very long in that situation but certainly an individual from getting from the back of the queue to the front you know I mean there are times when they wait for an hour in a corridor.

Reference 3 - 0.87% Coverage

a most common thing which tips us into the inefficiency that I talked about first and I mean in terms of the sort of set up what happens obviously the patients who would go to the medical assessment unit stay here and we very quickly fill up our beds and tip it to the out of control situation.

[<Internals\CS 2 Interviews\ 10 Nurse](#) - § 1 reference coded [1.47% Coverage]

Reference 1 - 1.47% Coverage

we try to work as effectively as possible sort of with them so either create the bed or if the patient is well enough we utilise our waiting area because we've got a clinic running with the waiting, next to the waiting area. So there is a senior registered nurse in there so there's the potential to continue to observe patients there and then bring them into the bed when we are able.

[<Internals\CS 2 Interviews\ 11 Nurse](#) - § 4 references coded [7.21% Coverage]

Reference 1 - 2.03% Coverage

So most people say to me you know medical outliers are not acceptable but yet they still seem to happen. So what is it that's allowing that to happen even though it's recognised as poor care?

Because the medical envelope of beds for whatever reason whether that's you know we've still not got our processes right or our entrances of exits right in terms of what we do with GPs, the bottom line is on that given day you do not have enough medical beds for those patients, therefore those patients have to go somewhere so they go to a surgical ward and become an outlier. So practically the other day that's a nuts and bolts fit we have too many patients for our complement of beds.

Reference 4 - 1.59% Coverage

what's it like on the medical wards itself trying to provide single rooms? All right well there aren't enough I mean we have a lot of data from infection control to know that we never have enough. I don't know where we sit as a comparable Trust with

other Trusts but I should think we are quite low actually on terms of our size and numbers and yes it's a constant challenge and that obviously creates more workload because you are trying then to move patients in and out of side wards constantly but we certainly don't have enough.

[<Internals\CS 2 Interviews\ 12 Doctor](#) - § 4 references coded [5.90% Coverage]

Reference 2 - 1.78% Coverage

I mean what, well I think the managers expect from us is that we go and see the patients and make a clinical decision and that's our side of the bargain. It's to be there as the senior clinician making a sensible decision for that patient. If the patient has to come in, has to be admitted, then that's the way it is. I have never had the managers saying well you can't do that we haven't got any beds you know. That would cause world war three I think.

Reference 3 - 1.45% Coverage

it's into the number of extra beds that have to be opened in order to as it were make that promise come true.

Yes. Yes, what's the option though? [LAUGHS] Yeah you know you can't turn away acutely ill patients so if it means you open extra beds and you have to cancel the routine work then that's what has to happen and I think everyone within this Trust agrees with that.

Reference 4 - 0.80% Coverage

So an average physician will do two ward rounds a week and their registrar will do a ward round a week and then the other days are SHO led and when we are in melt down there's often a physician in every day

[<Internals\CS 2 Interviews\13 Director](#) - § 5 references coded [14.31% Coverage]

Reference 1 - 1.74% Coverage

We have an additional capacity to cope with winter for which we had reserved half a million pounds at the beginning of the year and so that is essentially an additional ward and additional junior medical cover and some additional portering for both A&E and MAU.

Reference 5 - 1.61% Coverage

We don't have the diagnostics in the evening and that's what we need to you know get the patients out so that we've got patients staying overnight who could probably manage with less than a twelve hour stay if they came in earlier in the day.

[<Internals\CS 2 Interviews\ 14 Nurse](#) - § 4 references coded [5.33% Coverage]

Reference 1 - 1.11% Coverage

It's been extremely busy one of the reasons for the bed crisis is probable due to the fact that we've got an infection which happens to be on "G" Ward so G Ward's been closed on several occasions so for us it's not really affected us as much as it would have normally because we've not had to take any admissions because we've been closed to admissions due to the infection

Reference 2 - 1.19% Coverage

We've got nine beds empty at the moment on G but we have potentially got another three going home today, because patients can go to their own home but they can't be discharged to any other healthcare setting. So we are restricted to where they go but

we've got three going home so we will have eighteen patients today so, which gives us twelve empty beds.

Reference 4 - 1.49% Coverage

It depends what team of doctors it is really. If it's a team of doctors that are on your ward them it's usually quite, it's quite easy if it's a patient that's come up from medical assessment unit and they haven't yet been seen by the consultant's team then they are quite reluctant to see the patient because obviously they don't know the patient and out of hours it's very difficult to get, because the doctors have got so many admissions and things it's quite difficult to get the doctors to see the patient

[Internals\CS 2 Interviews\ 15 Doctor](#) - § 2 references coded [4.60% Coverage]

Reference 1 - 2.42% Coverage

in the day you will have possibly a registrar, possibly an SHO and a few house officers so it's probably about four of you as at five o'clock when it gets really busy clerking but the registrar often has to go on the ward round so there's probably three people clerking and admitting the patients until 10 and then after that you have two people covering the whole of medicine and there are doctors on for other specialities but up to, we haven't really developed a way of bringing them in.

Reference 2 - 2.17% Coverage

So I actually normally we just sort of cope. I think probably what you are meant to do well what they said meant to do but is call in a surgical SHO the obs and gynae SHO and the psyche SHO but these are doctors who are quite reluctant to come and help if they are in a speciality that's quiet and also they often haven't done medicine for quite a long time and sometimes I find if you get them to see somebody it just causes you more work

Case Study 3

Coded at CS 3: SWE; Patient Safety

Oral Witness B: The witness told the Inquiry that the unit (MAU) became known as "Beirut" throughout the hospital. She said that the low staffing levels made the unit dangerous from a safety perspective. It also meant that staff were not able to provide basic care. Patients were left without pain relief for long periods and in soiled wet bedding. Help with food was not given and patient buzzers were left ringing. Staff handovers were also not performed correctly.

Oral Witness C1: "My primary feeling, as is always the case when I am nursing, is would I want to be treated like this, or would I want a member of my family to be treated like this. And nine times out of ten the care was so appalling it would be no".

On CDU, patients, according to the Witness, "would be forgotten about". They would not be fed or given fluid or buzzers to ring for assistance. Medication was also routinely missed.

Oral Witness E1: believes that A&E was chronically understaffed during this entire period. The department could never triage patients due to understaffing and instead had

to rely on receptionists to judge the seriousness of cases and call for a nurse if a patient presented looking particularly unwell.

Coded at CS 3: SWE; Targets

Oral Witness C1: It was clear to the Witness that patients were moved in order to meet the four-hour target even when it was not in their best interest. The Witness said that patients, for example, were moved out of A&E when they were lying in soiled linen to prevent breaches. On another occasion, the Witness reported that a patient was discharged without being examined first by a doctor due to pressure from the Ward Sister. The patient returned to A&E the following morning where she died.

The Clinical Decision Unit (CDU) was used by A&E as a “dumping ground” the Witness said. Patients would be sent to CDU purely to prevent breaching the four-hour target.

Oral Witness G1: stated that the Trust was obsessed with achieving Foundation Trust (FT) status. In the Witness’s view, this obsession blinded the board to the problems with clinical care.

Oral Witness A11: spoke about the blame culture associated with targets. He recalled that if targets were not met you were required to explain this ‘failure’ to the Strategic Health Authority (SHA) and to report to them regularly until the problem was rectified.

Coded at CS 3: SWE; Finance

Oral Witness B1: The Medical Division in the hospital was required to make a saving of £581, 000. As a result, a reconfiguration of the Medical Division was implemented and Wards 10, 11 and 12 were amalgamated into Floor 2. A reduction of 20.4 whole-time equivalent posts (2 band 7s, 15.42 band 5 and 6s and 2.98 receptionists) was also proposed.

Oral Witness K1: the pressure to tackle the deficit in 2006 was not related to the hospital’s application for Foundation Trust (FT) status. The view at the time was that if the hospital was able to manage its finances and achieve FT status it would be given the freedom to allow it to manage its budget more effectively.

Oral Witness N11: At the Hospital Management Board (HMB) in March 2006 the Witness, recalled that the Chief Executive reported that the cost improvement programme had identified 4.6 million but that a shortfall of £5.4 millions remained that had to be tackled immediately.

Coded at CS 3: SWE; Workload

Oral Witness B1 The staffing situation on floor 2 was ‘desperate’, Witness B1 states, and she accepts this made the care provided unsafe.

Oral Witness H1: recalled that the standard of nursing care also declined as a result. There were fewer nurses on the surgical ward yet a greater proportion of high dependency patients to manage.

Oral Witness N11 (former Director of Finance): Trust did have information on establishment figures and on the number of people in post. Information was also available on the number of temporary staff who had been employed.

Oral Witness C (former Director of HR): Sickness levels were high, particularly among the Medical Division. Witness C said the Trust was also carrying around 200 vacancies, which they found it hard to reduce. She did not feel there was an adequate system for monitoring establishment figures because the payroll system was separate to the finance system. As a result, Witness C found it difficult to report precise figures about where vacancies existed within the hospital.

Coded at CS 3: RP; AD; Capacity

Oral Witness B1: Floor Two was originally intended to be a 64-bed floor. There were a number of contingency beds on the floor that were to be opened in the event of a bed crisis or if sufficient staffing for the beds was found. In reality, these additional beds were open for the majority of the time. Therefore, Witness B1 commented that the staff on floor 2 were expected to care for up to 82 patients as opposed to 64.

Oral Witness A11: One of the attractions for implementing the floors initiative was that it would allow you to ring-fence surgical beds that were being blocked by medical patients. This would therefore reduce the number of surgical patients who were deferred or delayed.

Appendix 6.3 – Results of interviewee questionnaire

Case Study 1 Questionnaire results

Total # placing issue as highest priority (*NED not able to answer)

Question #	Targets	Staffing	Safety	Finance	Totals	%	%	%	%
						Targets	Staffing	Safety	Finance
Board	4	0	15	9	28	14.3	0.0	53.6	32.1
Div Mgt	11	1	12	3	27*	39.3	3.6	42.9	10.7
Clinical	0	12	15	0	27*	0.0	42.9	53.6	0.0
Line Mgr	5	2	19	1	27*	17.9	7.1	67.9	3.6
You	4	5	17	1	27*	14.3	17.9	60.7	3.6
Totals	24	20	78	14	136	17.6	14.7	57.4	10.3

Total # placing issue as second priority

Question #	Targets	Staffing	Safety	Finance	Totals	%	%	%	%
						Targets	Staffing	Safety	Finance
Board	17	0	2	9	28	60.7	0.0	7.1	32.1
Div Mgt	11	2	5	9	27*	39.3	7.1	17.9	32.1
Clinical	0	15	11	1	27*	0.0	53.6	39.3	3.6
Line Mgr	9	8	3	7	27*	32.1	28.6	10.7	25.0
You	4	14	6	3	27*	14.3	50.0	21.4	10.7
Totals	41	39	27	29	136	30.1	28.7	19.9	21.3

Total # placing issue as third priority

Question #	Targets	Staffing	Safety	Finance	Totals	%	%	%	%
						Targets	Staffing	Safety	Finance
Board	7	4	9	8	28	25.0	14.3	32.1	28.6
Div Mgt	4	6	7	10	27*	14.3	21.4	25.0	35.7
Clinical	19	0	0	8	27*	67.9	0.0	0.0	28.6
Line Mgr	8	3	5	11	27*	28.6	10.7	17.9	39.3
You	13	3	3	8	27*	46.4	10.7	10.7	28.6
Totals	51	16	24	45	136	37.5	11.8	17.6	33.1

Total # placing issue as fourth priority

Question #	Targets	Staffing	Safety	Finance	Totals	%	%	%	%
						Targets	Staffing	Safety	Finance
Board	0	24	2	2	28	0.0	85.7	7.1	7.1
Div Mgt	1	18	3	5	27*	3.6	64.3	10.7	17.9
Clinical	8	0	1	18	27*	28.6	0.0	3.6	64.3
Line Mgr	4	15	0	8	27*	14.3	53.6	0.0	28.6
You	5	6	1	15	27*	17.9	21.4	3.6	53.6
Totals	18	63	7	48	136	13.2	46.3	5.1	35.3




Case Study 2 Questionnaire results

<i>No ranked 1 (highest)</i>	Targets	Staffing	Safety	Finance	
Trust Board	3	0	7	5	15
Mgt Teams	2	0	12	1	15
Clinical teams	0	4	10	0	14 * NED not able to answer
Percentages					
Trust Board	20	0	47	33	
Mgt Teams	13	0	80	7	
Clinical teams	0	29	71	0	
<i>No ranked 2</i>	Targets	Staffing	Safety	Finance	
Trust Board	5	0	3	7	15
Mgt Teams	4	6	1	4	15
Clinical teams	0	10	4	0	14 * NED not able to answer
Percentages					
Trust Board	33	0	20	47	
Mgt Teams	27	40	7	27	
Clinical teams	0	71	29	0	
<i>No ranked 3</i>	Targets	Staffing	Safety	Finance	
Trust Board	6	1	5	3	15
Mgt Teams	6	4	1	4	15
Clinical teams	11	0	0	3	14 * NED not able to answer
Percentages					
Trust Board	40	7	33	20	
Mgt Teams	40	27	7	27	
Clinical teams	79	0	0	21	
<i>No ranked 4</i>	Targets	Staffing	Safety	Finance	
Trust Board	1	14	0	0	15
Mgt Teams	3	4	1	7	15
Clinical teams	3	0	0	11	14 * NED not able to answer
Percentages					
Trust Board	7	93	0	0	
Mgt Teams	20	27	7	47	
Clinical teams	21	0	0	79	

Targets = Achieving targets
 Staffing = Adequate staffing
 Safety = Patient safety
 Finance = Achieving financial results

Appendix 7.1 - Examples of Daily Bed Information

Case Study 1

B1		11/12/2008										
A	B	C	D	E	F	G	H	I	J	K	L	
1	Daily Forecast Thursday 11 December 2008 AM											
2												
3												
4	Trust Barometer		Additional Relevant Information									
5			<p>CDU OPEN & FULL CAPENER BEING USED AS AN INPATIENT AREA FOR MEDICINE</p> <p>WE ARE LOOKING FOR ANY ASSISTANCE WITH STAFFING THE EXTRA AREAS FROM MIDDAY TODAY, PLEASE CONTACT THE SITE OFFICE IF YOU ARE ABLE TO HELP.</p>									
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17		Predicted Admissions			Discharges							
18	Directorate	Emergency	Inpatient	Elective Daycase	From Admission Areas	Beds Available at 08:00	Required Ward Discharges	Identified Ward Discharges	Shortfall at midday			
19	Medicine	42	11	3	13	0	40		40			
20	Surgery	18	23	14	5	-4	40	22	18			
21	Trauma and Orthopaedics	4	13	8	5	4	8		8			
22	Cancer Services	4	1	64	0		5		5			
23	Gynaecology	5	9	9	0	7	7	6	1			
24	Paediatrics	9	8	0	0	22	-5	4	-9			
25	Total	82	65	98	23	29	95	32	63			
26	EMU beds at 08:00								0			
27	SAU beds at 08:00								-4			
28	Number of Wards Closed by IC								11			
29	Total Patients Ready on Community List								19			
30	Patients Ready but on closed areas								5			
31	Predicted Community Discharges Today								6			
32												
33	Pressure Indicators											
34												
35	Medical outliers								56			
36	15 plus day stayers								139			
37	Patients on and awaiting addition to Trauma list								13			
38	Patients on and awaiting addition to CEPOD list								5			
39	Patients awaiting Acute Stroke Unit								1			
40	Extra daycase capacity used / CDU open								19			
41												
42	Emergency Department											
43												
44	Day	Breaches	Attendances									
45	Monday	25	297									
46	Tuesday	6	219									
47	Wednesday	12	245									
48	Thursday											
49	Friday											
50	Saturday											
51	Sunday											
52	Position	94.3%										
53												
54	Intensive Care Unit											
55												
56	Number of level 3 patients								5			
57	Number of level 2 patients								5			
<p>Summary Adm Area Discharges Today's Predicted Disch Tomorrow's Predicted Disch</p>												

Case Study 1

Microsoft Excel - 05-01-2009_AM.xls

File Edit View Insert Format Tools Data Window Help

D23

A B C D E F G H I J K


1

2 **Daily Forecast Monday 05 January 2009 AM**

3

4 **Trust Barometer** **Additional Relevant Information**

5

6  Severe

7

8

9 Moderate

10

11 Good

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

NO INFORMATION RECEIVED FROM PAEDS

Exe & Tavy wards currently being used in Medical capacity
1 Bay on Torridge being used in Medical capacity
ORGANISATION UNDER SEVERE PRESSURE TODAY

Directorate	Predicted Admissions			Discharges From Admission Areas	Beds Available at 08:00	Required Ward Discharges	Identified Ward Discharges	Shortfall at midday
	Emergency	Inpatient	Daycase					
Medicine	42	7	13	11	0	38		38
Surgery	16	29	11	4	1	40		40
Trauma and Orthopaedics	6	11	5	4	6	7		7
Cancer Services	4	4	57	0	2	6		6
Gynaecology	8	6	4	0	6	8	5	3
Paediatrics	12			0		12		12
Total	88	57	90	19	15	111	5	106

EMU beds at 08:00	0
SAU beds at 08:00	0
Number of Wards Closed by IC	1
Total Patients Ready on Community List	40
Patients Ready but on closed areas	0
Predicted Community Discharges Today	6

Pressure Indicators

Medical outliers	113
15 plus day stayers	140
Patients on and awaiting addition to Trauma list	4
Patients on and awaiting addition to CEPOD list	
Patients awaiting Acute Stroke Unit	4
Extra daycase capacity used / CDU open	24

Emergency Department

Day	Breaches	Attendances
Monday	1	235
Tuesday	3	186
Wednesday	9	179
Thursday	3	188
Friday	24	221
Saturday	8	213
Sunday	18	200
Position	95.4%	

Intensive Care Unit

Number of level 3 patients	4
Number of level 2 patients	7

Summary Adm Area Discharges Today's Predicted Disch Tomorrow's Predicted Disch

Ready

Case Study 2

Microsoft Excel - Daily_Bed_Status 22.1.10(1).xls

File Edit View Insert Format Tools Data Window Help

E4 Page down for the comments box

	A	B	C	D	E	F	G	H	I
1	Bed Status	21/01/2010	19:19						
2									
3	Acute admissions in the last 24 hours (8am - 8am)								
4	Medicine	32							
5	Surgery	2							
6	Orthopaedics	0							
7	Paediatrics	10							
8	Gynae	2							
10	Empty beds after NBMs placed @1pm				<u>Remaining</u>	<u>Forecast</u>			
11	Medicine	3			<u>Predicted</u>	<u>Acute</u>	<u>Planned</u>		
12	MAU	-7			<u>Discharges</u>	<u>Admissions</u>	<u>Admissions</u>		
13	Surgery - elective (KGV)	-2			0	19	0		
14	Surgery - emergency (FAU)	-2			0	6	1		
15	Orthopaedics - emergency (Cape)	4			0	2	0		
16	Orthopaedics - elective (Lundy)	2			7		9		
17	Paediatrics	0			0	4	0		
18	Gynae	-6			4	0	4		
20	Number of available adult beds:	5			Action required to manage bed capacity				
21	Number of forecast adult admissions:	43			38 more beds required				
23	Total NBMs	0							

Microsoft Excel - Daily_Bed_Status 22.1.10(1).xls

File Edit View Insert Format Tools Data Window Help

E4 Page down for the comments box

Dear all,

the bed situation is in the black.

Petter at 19 patients.

Overnight day surg to be open overnight to start with 4 patients in there.

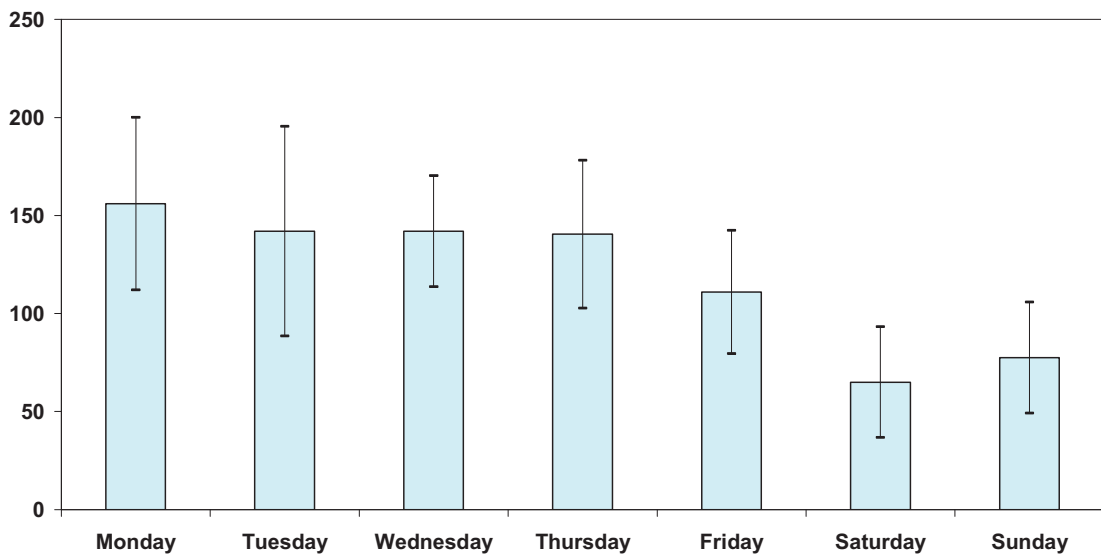
ENT list for day surg as cancelled.

To staff day surg very difficult, as all wards full and busy.

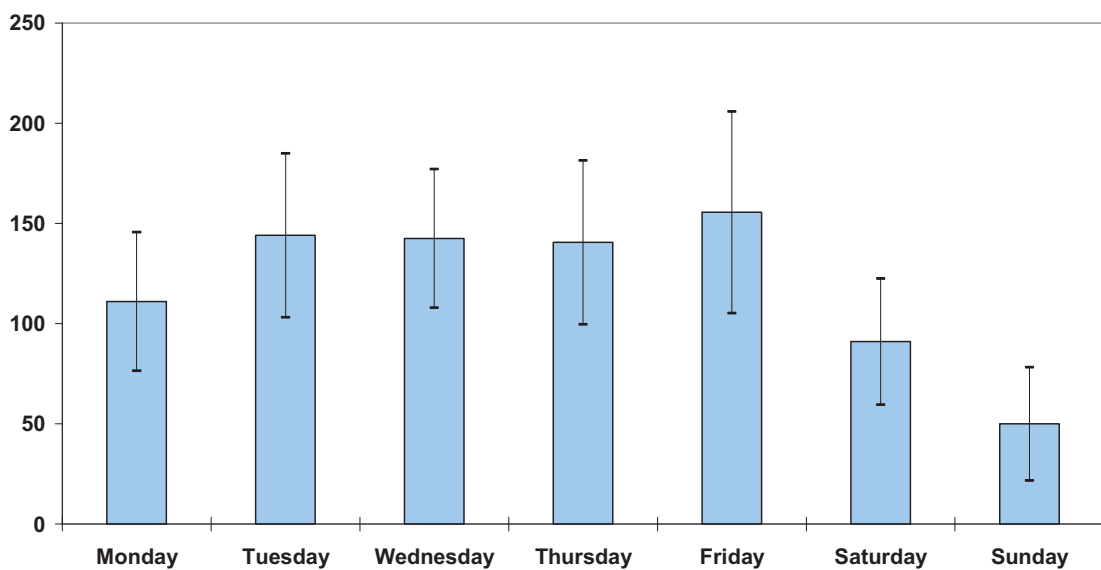
Thank Ellen

Appendix 8.1 – Case Study 1 – admission and discharge graphs 2006/07 (source CS 1 xls 1.11)

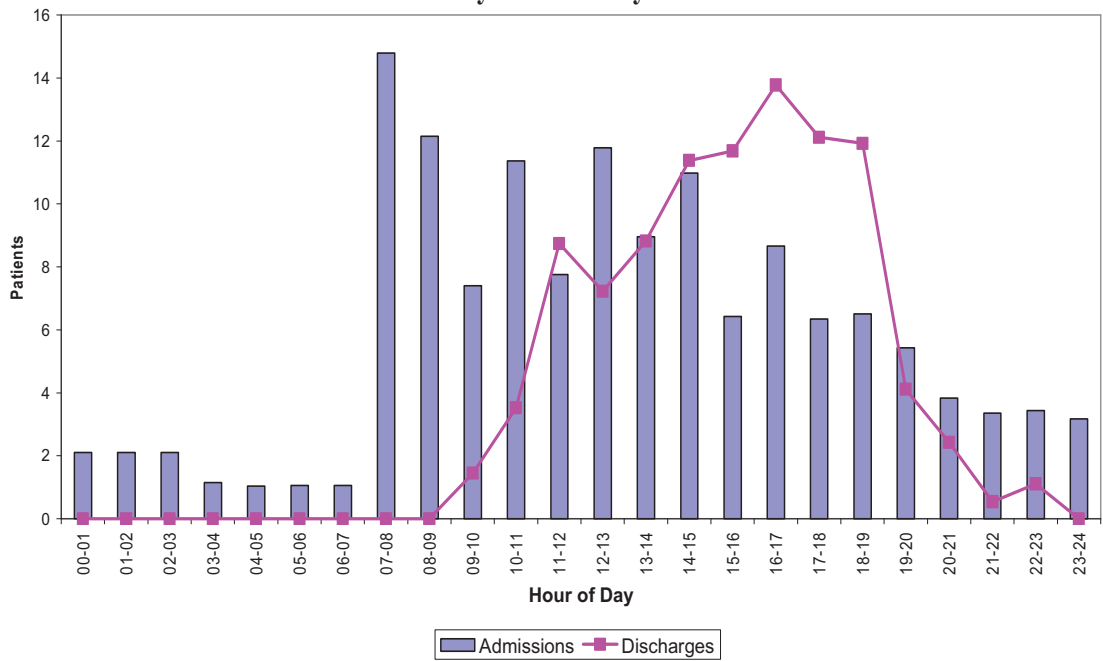
Case Study 1- Elective and Non-Elective Admissions by Day of Week 2006/07
Median, Upper and Lower Control Limits



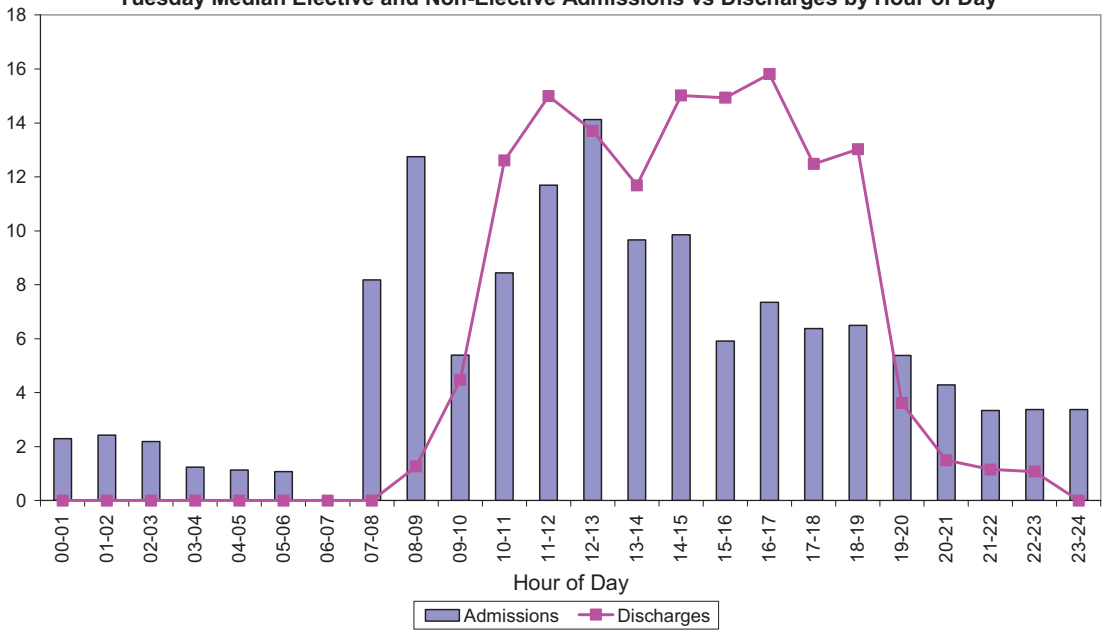
Case Study 1 - Elective and Non-Elective Discharges by Day of Week 2006/07
Median, Upper and Lower Control Limits



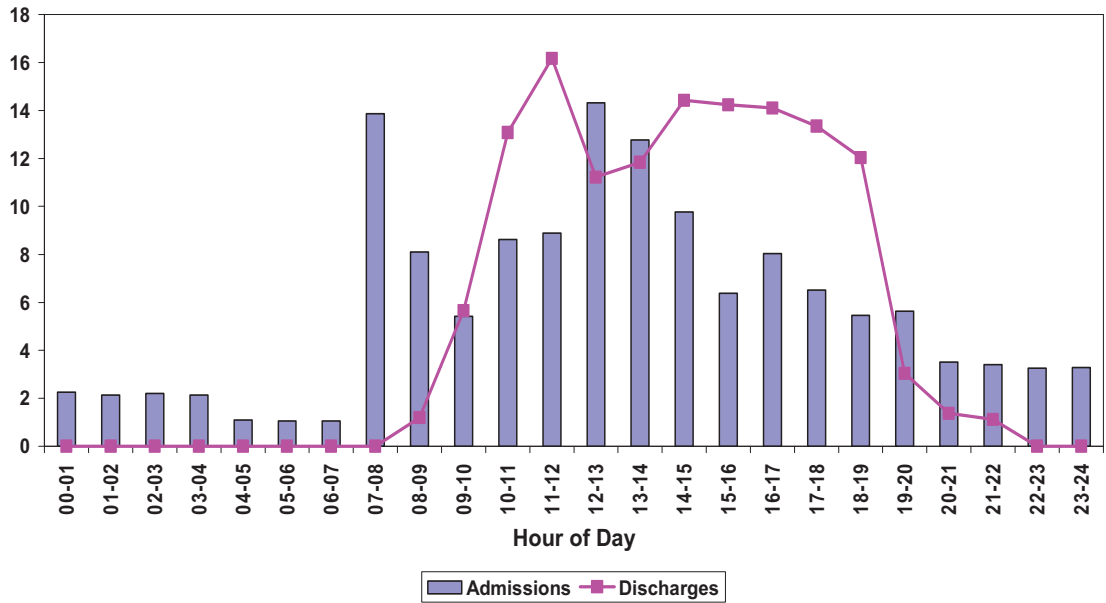
CS 1 Monday Median Elective and Emergency Admissions vs Discharges by Hour of Day



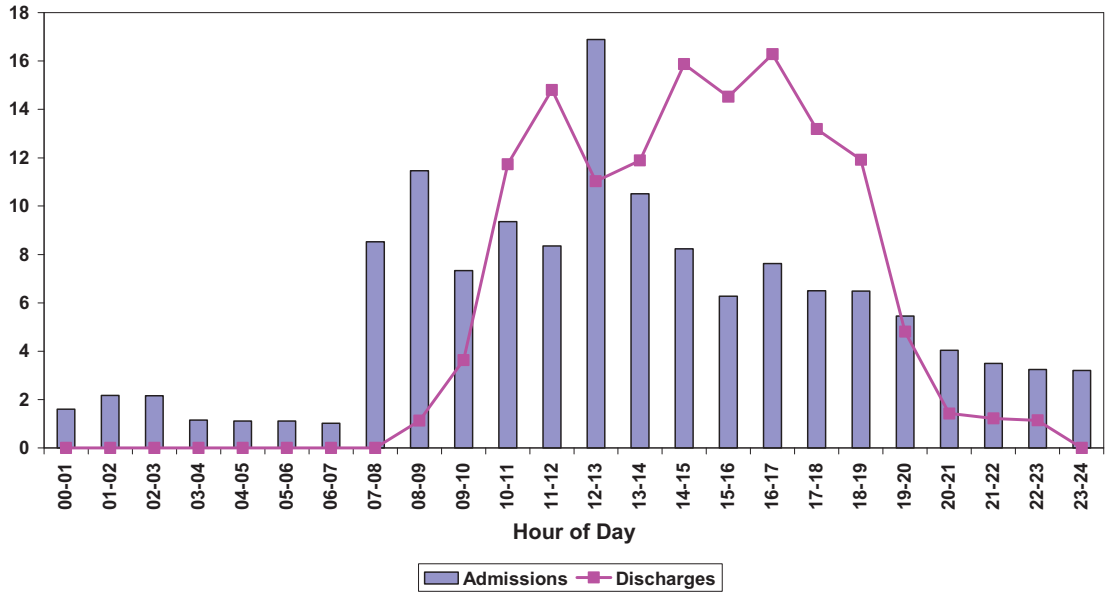
CS 1 Tuesday Median Elective and Non-Elective Admissions vs Discharges by Hour of Day



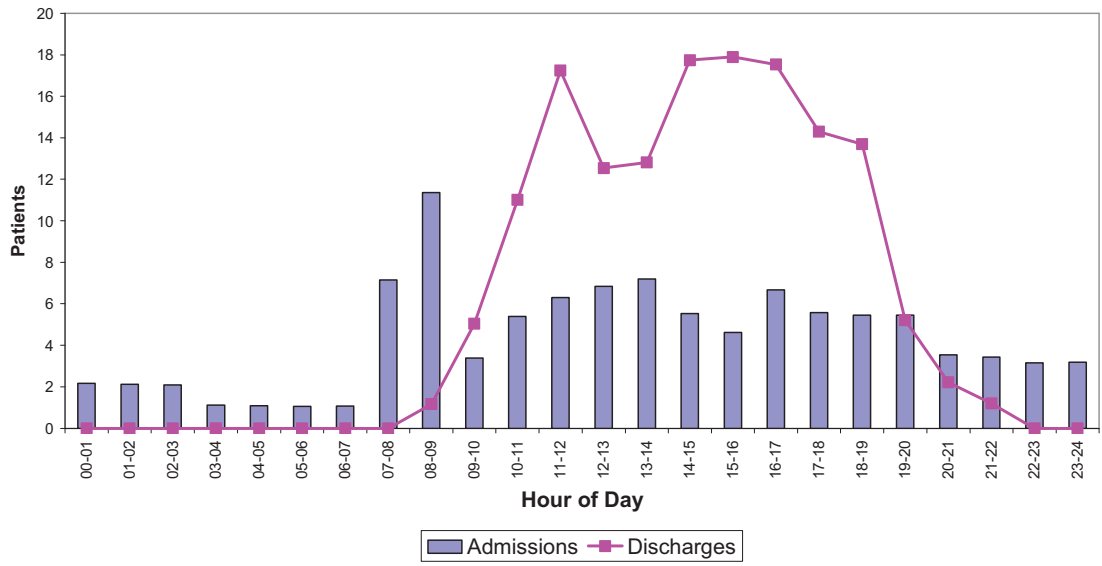
CS 1
 Wednesday Median Elective and Non-Elective Admissions vs Discharges by Hour of Day



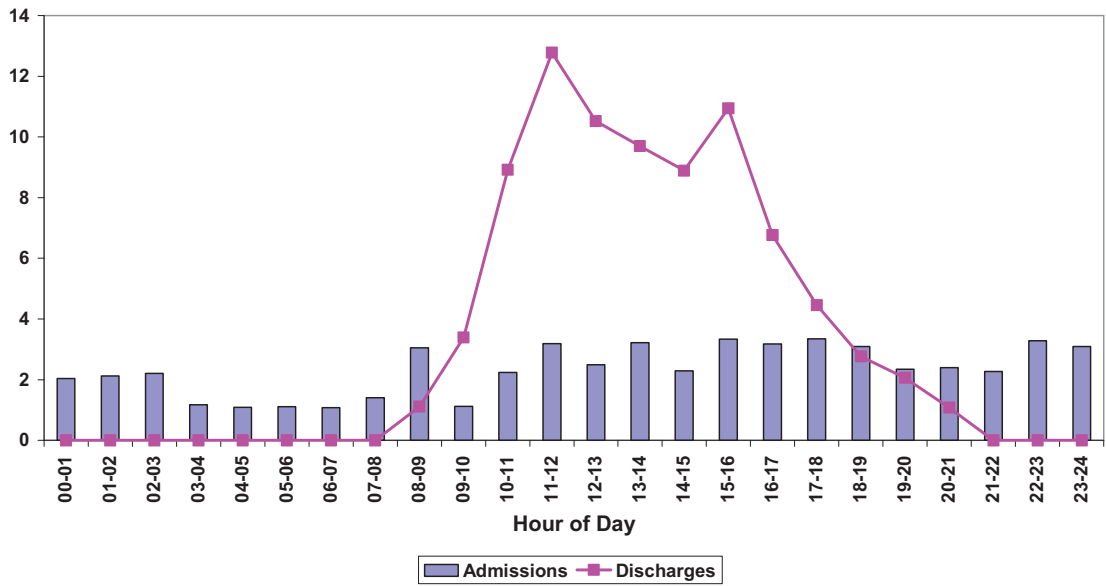
CS 1
 Thursday Median Elective and Non-Elective Admissions vs Discharges by Hour of Day



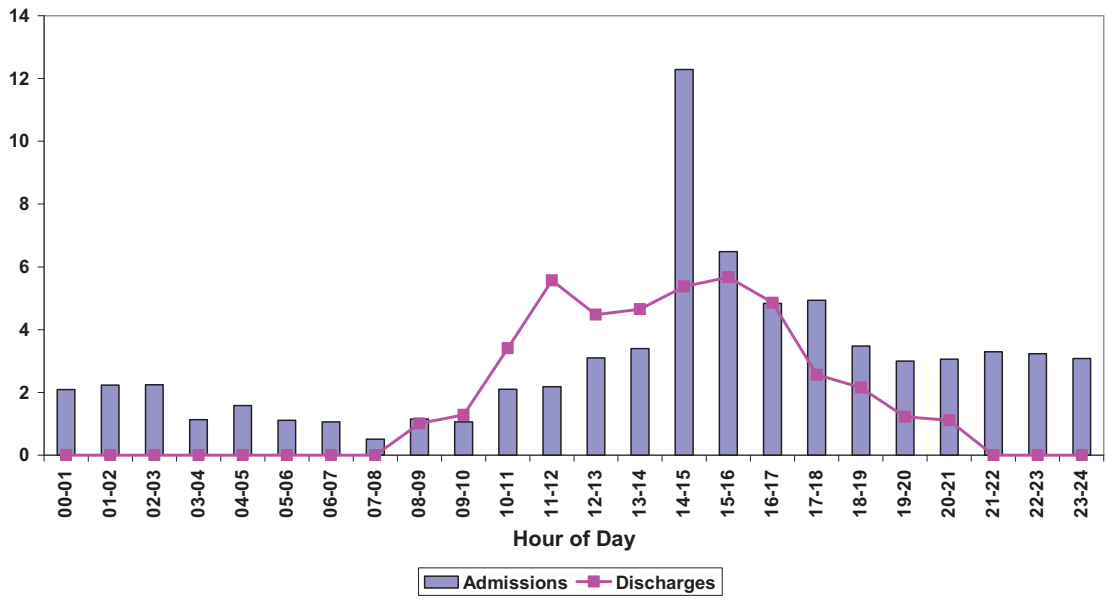
CS 1 Friday Median Elective and Emergency Admissions vs Discharges by Hour of Day



CS 1 Saturday Median Elective and Non-Elective Admissions vs Discharges by Hour of Day



CS 1
 Sunday Median Elective and Non-Elective Admissions vs Discharges by Hour of Day



Publication 1

Williams M.D, Smart P.A. (2010) Patient Safety: a casualty of target success?
International Journal of Public Sector Management, 23, 416-430.



Patient safety: a casualty of target success?

Mike Dermot Williams and Andi Smart
University of Exeter Business School, Exeter, UK

416

Received 30 September 2008
Revised 4 February 2009
Accepted 3 June 2009

Abstract

Purpose – This paper aims to develop a conceptual resilience-based model that takes account of the competing success factors of patient safety, finance, improvement targets and staff workload in NHS hospitals in the UK.

Design/methodology/approach – A safe working envelope model was developed from the literature and adapted for use in the NHS. The proposition that finance and targets receive greater management attention was then tested by a pilot study using content analysis of risk management documents of four NHS hospitals.

Findings – The need to succeed on finance and targets received greater attention in the risk management documents than patient safety and staff workload.

Research limitations/implications – This is a pilot study only, using content analysis of risk management documents from four hospitals to see whether the model developed from the literature warrants further study.

Practical implications – Using the proposed safe working model will allow the setting and monitoring of failure and marginal boundaries and make more explicit the pressures from the competing success factors in public sector hospitals in the UK.

Originality/value – The development of the conceptual model using ideas from resilience engineering and applying them to NHS hospital management provides a policy and practical approach to improving patient safety.

Keywords Patients, Safety, Health care, Risk analysis, Targets, National Health Service

Paper type Conceptual paper

What counts as success in public sector health care? There is no shortage of performance measures seeking to provide the answer. From a patient safety perspective, there is an imbalance with the current emphasis of those performance measures on finance and targets, within the National Health Service (NHS) in England. Keeping patients safe from unintended harm is a significant issue and fundamental for the both the patient and the delivery system providing treatment. Within the National Health Service (NHS), as in other developed countries, there is a poor record for patient safety (DoH, 2000).

The scale of the problem is significant. Adverse events have been described as “incidents in which a patient is unintentionally harmed by medical treatment” (Vincent *et al.*, 1998). Vincent *et al.* (2001), in a retrospective case note study found 10.8 per cent of patients suffering adverse events in two NHS hospitals, almost half of which were judged preventable. Sari *et al.* (2007) in a similar study of a large NHS teaching hospital, concluded that 8.7 per cent of admissions had at least one adverse event, of which 31 per cent were judged preventable. The study showed that 15 per cent of the adverse events led to impairment or disability that lasted more than six months and 10 per cent contributed to the patient’s death. The result for patients was a prolonged



period of treatment leading to an increased mean length of stay by 8 days. In a more recent prospective study focusing on an admission ward, the rate of adverse events was found to be 11.4 per cent and potential adverse event or near miss to be 14.8 per cent (Catchpole *et al.*, 2008).

The seriousness of the issue was recognised by the Chief Medical Officer within the Department of Health (DoH) in a landmark report, “An Organisation with a Memory” (DoH, 2000). A consequence of this report was the establishment of the National Patient Safety Agency (NPSA) which developed a national reporting scheme for errors and near misses. Following an Audit Office Report on the lack of progress made on patient safety (National Audit Office, 2005), the DoH issued a further report “Safety First” (DoH, 2006a). As part of the drive to improve patient safety the DoH also set out “core standards” of safety within the “Standards for Better Health” (DoH, 2006b). These are standards that all health care providers must meet. Compliance with all the standards is monitored annually by an independent body, the Healthcare Commission.

Given the serious nature of the problem and the considerable policy initiatives it would be expected that significant progress would have been made in measuring and reducing the level of adverse events within NHS hospitals. However, as will be shown, the best intentions of making “safety first” has become a casualty of other more pressing success factors, such as waiting times and financial balance.

Competing priorities in public sector services and particularly in health care, is nothing new. What is needed is a way to conceptualise those priorities and understand how the resulting competing success criteria are described and measured. Such a conceptualisation needs to take account of the complex interactions of priorities within the system of public healthcare in such a way as to give sufficient weight to safety for patients. This paper sets out a “system resilience” approach to understanding patient safety within the wider context of competing priorities in NHS hospitals.

Literature

Healthcare systems have multiple professional and stakeholder groups, low reliability processes (Resar, 2006), macro and micro system interactions (Mohr *et al.* 2004), fragmented leadership, diffuse power and multiple goals (Lozeau *et al.*, 2002). Reason (1997) argues that to achieve improved safety, attention has to be paid to the sharp and blunt end of the system. Those at the sharp end are those people delivering patient treatment. They are human and are likely to make mistakes (Reason, 1990). Equally those managers with responsibility for the organisation make decisions at the blunt end (removed from direct delivery), which create the conditions where mistakes can be made more easily by staff at the sharp end. Sheridan (2008), reviewed the ideas behind the traditional approaches to risk and found considerable weaknesses in their ability to consider future safety issues. He suggests that the resilience engineering paradigm has strengths in viewing safety from an organisational process perspective. Resilience in this context is about anticipating, mitigating and preparing to recover from unsafe events (Sheridan, 2008).

Woods and Hollnegal (2006) suggest that resilience engineering is an approach to safety that “focuses on how to help people cope with complexity under pressure to achieve success”. It is not about just counting “error” and then acting to reduce that count but rather putting safety as a core value in an organisation. Woods (2006) in the context of healthcare describes “resilience” as meaning “a work system’s ability to

buffer, adapt to, absorb and prevent adverse patient outcomes in the face of disruption”.

Resilience is a proactive approach to safety which, if successful, means that the system can keep operating in what resilience engineering theorists describe as an “operating envelope” within which the system is designed to function (Rasmussen, 1997; Woods, 2006). Rasmussen (1997) describes three interacting boundaries to the safe working envelope; the boundary of economic failure, the boundary of unacceptable workload and the safety boundary of unacceptable performance. The model suggests that there is a marginal boundary or zone inside those boundaries which if breached would create the conditions for failure (Figure 1). The “operating point” can be described as where the competing pressures reach a theoretical position of equilibrium.

Woods (2006), states that unanticipated problems occur either because the three boundary model is limited/wrong or the environment changes which bring pressures to bear on the level of functioning in the system. The pressures from the two boundaries of economic failure and workload may push the system’s operations, known as the “operating point”, into the marginal zone either through a process of “drift” or through rapid change that may then result in an accident (Figure 2). This could also be described as there being incremental or step change in the location of the “operating point”. Decisions made by any number of stakeholders in a public service can create pressure on the operating point, even when those decisions are made at a level far removed from direct patient care (Barber, 2008). Within the NHS this can mean that the pressures which are “in focus” are the management of finance and targets.

Decision makers live in a world of conflicting goals with many consequential dilemmas. To choose one side of a dilemma (e.g. production to achieve targets) can create a hidden condition in the system on the other side of the dilemma (e.g. safety) (Reason, 1997). Those making such decisions do not always examine the consequence, or sacrifice, of their decisions before making them as they maybe over focused on achieving certain success criteria. Nor is it often easy to understand what the result might be even if time was given to such consideration. The result may be that nothing

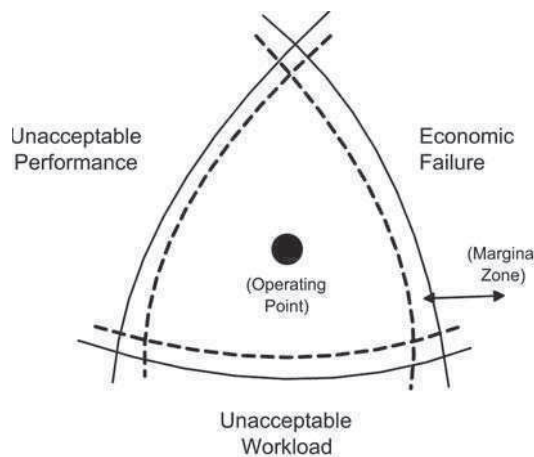
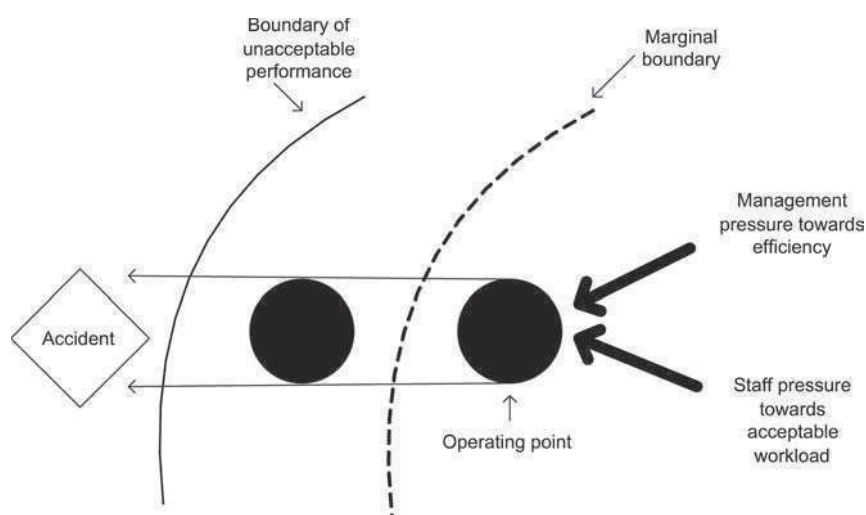


Figure 1.
Rasmussen’s modified
three-boundary safe
working envelope

Source: Adapted from Rasmussen (1997)



Source: Adapted from Miller and Xiao (2006)

Figure 2.
Pressure on the operating point

untoward happened when production is favoured against safety. However, the system may well have moved closer to the boundary of unsafe operations where a new “normal” way of working is established. This unobserved movement of the operating point towards the boundary is a concept that has been termed “drift to danger” (Rasmussen, 1997).

Vaughan (1996) shows that normalisation of deviance is key issue for safety. Her study of the Challenger Space Shuttle disaster illustrated that the system within which people work can produce a culture where, through small incremental steps, new situations are seen as “normal”. This is described as the “native view”. Outsiders looking at the situation are more likely to see the situation as deviant, not acceptable and therefore potentially dangerous. Waring *et al.* (2007) point to types of behaviour by health care workers that illustrate “taken for granted assumptions about clinical risk...” Such behaviours “normalize risk” and as such mean that risk of harm to patients is not addressed. Waring (2005) suggests that medical staff regard error as “inevitable” which can lead to errors being seen as “normal”. These errors may not be reported as incidents and therefore do not provide the basis for improvements in safety. Normalization in the context of the safe working envelope can be regarded as a “drift to danger” of the operating point. Equally, normalization can be conceived of as shifting the safety failure boundary to a new but more dangerous position.

The Healthcare Commission (2008) has looked at lessons learnt from undertaking fourteen investigations into failures in NHS services. It concluded that some boards of NHS trusts are “particularly vulnerable to being consumed by the business of healthcare, in the form of mergers, reconfiguration of services, financial deficits and targets”. They were surprised by the extent to which organisations did not have adequate systems in place to know about potential quality problems. The report argues that patient safety should not be compromised by other objectives.

Cook and Rasmussen (2005), use Rasmussen’s safe working envelope model to examine safety problems for hospitals that are overfull. They suggest that it is normal

for healthcare systems to work at the limit of their capacity and for Hirschhorn’s “law of stretched systems” to apply: “every system always operates at its capacity. As soon as there is some improvement, some new technology, we stretch it . . .”. Stretch in this context means that the system is thought to be capable of that new level of performance. The underlying strain on the human aspects of the system is not always taken into account.

When systems are under resource and/or performance pressures, the benefits of change are taken in increased productivity or efficiency. The efficiency pressure on the operating point is such that it is moved closer to the boundary of unacceptable performance – namely safety failure. Woods and Cook (2002), describe this as systems moving back to the “edge of the performance envelope”. Studies have shown an association, but not proven causation, between hospitals that have high occupancy rates and an increase in mortality and adverse events (Richardson, 2006; Sprivilis *et al.*, 2006; Weissman *et al.*, 2007).

Miller and Xiao (2007) have built on the ideas of Cook and Rasmussen (2005). They recognise that when “bed gridlock” occurs patients wait longer at different points in the system which increases the chances of adverse events such as delayed diagnosis. This in turn is likely to increase the length of stay, thereby increasing the complexity in terms of the number of patients to be treated within the organisation and the bed capacity problem. Miller and Xiao studied the strategies used by staff to respond to high patient demand in a large trauma unit. They suggest that staff undertook “compensating actions” such as flexible rotas and schedules, to keep the operating point close to the marginal boundary of safety failure. Their theory is that “backup behaviours that make up the marginal zone may act as measures of the system’s resilience”. In other words, the marginal zone is the area in which compensation actions occur. Such actions constitute what has been described as “resilience”, and they argue, as such can be measured (Figure 3).

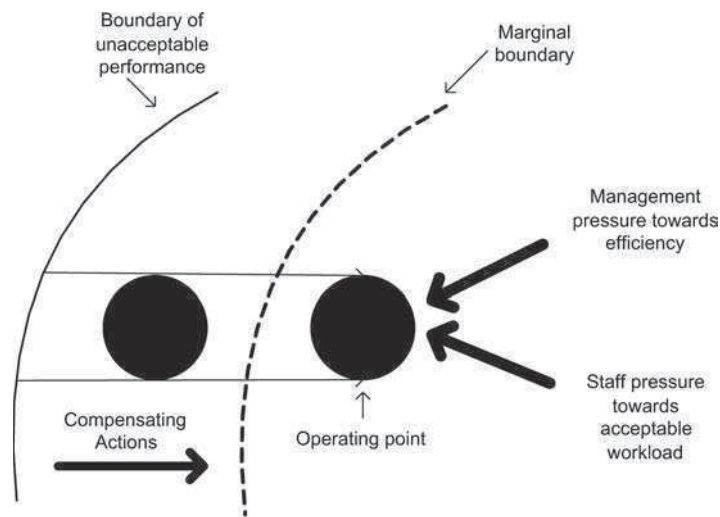


Figure 3. Compensating action in the marginal zone return the operating point to within the safe working envelope

Source: Adapted from Miller and Xiao (2007)

Proposed descriptive model of resilience for NHS hospitals

Building on the work by Cook and Rasmussen (2005) and Miller and Xiao (2007), we develop a descriptive model to explain the resilience of NHS healthcare systems in their wider context by proposing four boundaries of a “safe working envelope”. The development of the model seeks to reflect the additional aspect of the politically driven improvement targets (Barber, 2008) which combine with the pressures described by Rasmussen (1997). The four are the boundaries of financial failure, target failure, unacceptable working conditions, and failure of safety (See Figure 4). The financial failure boundary sets the limit of the organisation’s budget. To move outside that boundary means a deficit situation has arisen. The target failure boundary covers the numerous waiting time and other improvement targets set nationally and locally. Failure to meet any of those takes the operating point outside the boundary. The working conditions boundary relates to the pressure on staff to work above their contracted hours or at an unsustainable pace. Defining and measuring that boundary is far less clear than the first two. The failure of safety is when something goes wrong for a patient or groups of patients. That boundary is not always easy to clearly identify.

This descriptive model seeks to explain the constantly changing pressures that apply to the concept of a safe working envelope for NHS hospitals. It is not exhaustive in showing the external pressures but seeks to illustrate that there are a number of stakeholders who sometimes place conflicting goals on NHS hospitals. For example, to achieve high quality and safe care means providing treatment for patients in the most

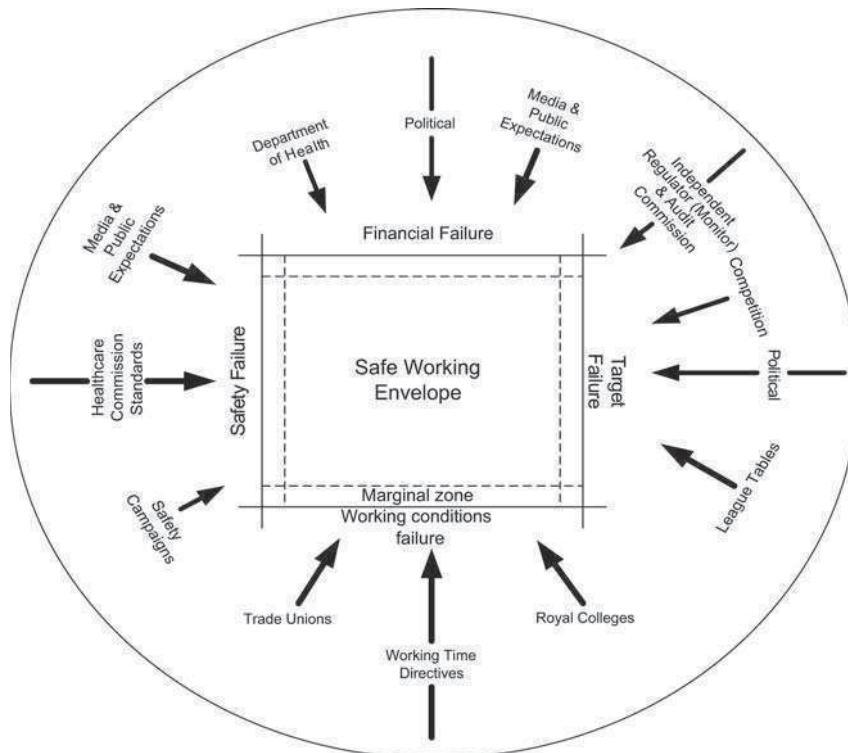


Figure 4.

A descriptive model of a safe working envelope for an NHS hospital

appropriate setting in a hospital. The waiting time targets can mean moving medical patients within the hospital, including in the early hours of the morning, to surgical wards in order to create suitable vacant beds for emergency medical patients to be admitted to within the four hour target from the Accident and Emergency department.

Within the wider context there is a strong political and managerial requirement to achieve financial balance or better whilst at the same time reducing waiting times to meet the targets. At the same time there are nationally negotiated staff contracts that specify the working arrangements for the staff which limits the availability of key groups, such as junior doctors.

For NHS hospitals to manage their risk in a way that takes account of the dynamic pressures that occur in complex organisations they need to establish a means of knowing where their operating point is located in relation to the boundaries of the safe working envelope. One way of doing this is to make the pressures influencing the operating point and the subsequent responses from managers and staff explicit. It is also possible to know when certain boundaries have been breached. For example, data for when the boundary of safe working is breached is obtained when errors occur (Woods and Cook, 2006). Most NHS hospitals have a system of reviewing and learning from incident reports, yet this has limitations. Many incidents go unreported and others are unrecognised as failures (Olsen *et al.*, 2007). Therefore, just examining incident reports is not a reliable means of knowing when the operating point has breached the safety boundary. Resilience, being a proactive concept of safety, means that other more predictive measures are required. Theoretically if we can measure what compensating actions are being undertaken in the marginal zone, then we can build a picture of where the operating point is located.

Compensating actions as a measurable indicator of resilience

From the work of Miller and Xiao (2007) we propose examining the nature and extent of “compensating actions” at different levels in the system as a means to find where the operating point is located. For example, we could suggest that when there were sufficient beds empty in the medical unit at the beginning of the day to accommodate the expected maximum number of medical emergencies and there were no medical patients in surgical beds (outliers), then the operating point would be regarded as being well within the safe working envelope. If this situation was normal each week then we could suggest that the operating point was also stable inside the envelope. If, however, at the beginning of the day there were no empty beds but sufficient expected discharges to accommodate the maximum number of expected medical emergencies and no medical outliers; then we could suggest that the operating point was close to breaching the marginal boundary.

With the process of normalisation there then may be some dispute as to when the operating point crosses the marginal boundary. In many hospitals a new “norm” has been observed where it is common to have medical outliers. Therefore staff would consider the operating point in those circumstances to be within the safe working envelope as the situation is “normal” (Waring *et al.*, 2007). However, in theoretical terms once compensating actions take place, such as putting medical patients in surgical beds (outliers) or “stacking” at home the emergency admissions being requested by General Practitioner (GP) as a means to cope with demand (Proudlove *et al.*, 2003), then

it could be suggested that the marginal boundary had been crossed using GPs as buffer capacity.

We could suggest that there is a breach of the safety boundary when “bed gridlock” occurs, all beds are full and admissions are waiting in corridors or ambulances with the consequent increased tight coupling across the system. In such situations the impact on staff working at the sharp end is such that the patients care is more likely to be compromised (Cook and Rasmussen, 2005). Compensating actions can be formal and widely communicated and therefore be monitored by management through a traditional performance management route. However, there are also numerous informal compensating actions that staff take, often known as “work arounds” which are not measured or monitored and could be hiding the movement of the operating point into the marginal safety zone.

Boundary setting and monitoring as a method of quantifying resilience

Recognising the competing and dynamic priorities that occur in healthcare systems (Woods, 2006) and the potential for the operating point to “drift to danger” (Rasmussen, 1997), it is necessary for managers to be able to describe what constitutes the marginal and failure boundaries. From the literature on stretched systems (Cook and Rasmussen, 2005, Woods and Cook, 2002), the competing priorities (Woods, 2006) and the NHS Operating Framework 2008-2011, (DoH, 2007), we can make the proposition that the boundaries of financial failure and target failure in NHS hospitals would be more clearly described and monitored than the other boundaries. To clarify this proposition we undertook a pilot study that involved content analysis of key risk management documents from four NHS hospitals.

NHS hospitals have risk management systems set out by the Department of Health (DoH, 2003). They are designed to allow hospitals to identify their key objectives and the risks associated with achieving them. Within the risk management system there is an “Assurance Framework”. This is defined as “a structure within which boards identify the principal risk to the organization meeting its principal objectives and map both the key controls in place to manage them and also how they have gained sufficient assurance about their effectiveness” (DoH, 2003). The Assurance Framework is then monitored at hospital board level on a regular basis through the year.

Content analysis of risk assurance frameworks of NHS hospitals

Method

The pilot sample was chosen by using the Dr Foster Report on NHS hospital mortality rates (Dr Foster, 2007). A hospital was chosen from each of the reported high, medium or low mortality rate categories. One other hospital was chosen that had featured in the Healthcare Commission Report on investigations into mortality (Healthcare Commission, 2008). Their Assurance Frameworks were accessed via the internet.

Data collection

A protocol was devised using the four boundary model. For each boundary a number of key words were identified as coding themes (see Table I). These codes were derived deductively from the literature on safe working envelopes, patient safety campaign interventions (Institute for Healthcare Improvement, 2008), the NHS Operating Framework 2008-2011, (DoH, 2007). Congruent with a grounded theory approach

Table I.
Assurance framework
content analysis of failure
categories

N	Hosp. A	Hosp. B	Hosp. C	Hosp. D	N	A	B	C	D
<i>Financial failure</i>									
1	29	8	5	14	1	10	11	72	37
2	10	2	3	3	2	0	0	2	0
3	1	2	0	0	3	5	14	31	8
4	3	0	1	3	4	0	0	7	0
5	5	4	2	4	5	2	2	13	0
6	0	0	0	0	6	12	17	0	18
7	7	0	0	3	7	4	2	4	0
8	12	0	0	13	8	2	12	27	2
9	1	0	0	0					
10	1	0	0	0					
11	11	1	1	5					
12	4	0	0	0					
Total	84	17	12	45		35	58	156	65
Mean	7.0	1.42	1.0	3.75		4.38	7.25	19.5	8.13
<i>Unacceptable workload</i>									
1	5	0	4	8	1	9	5	0	5
2	0	0	0	0	2	2	5	2	6
3	5	0	0	0	3	0	0	0	1
4	0	0	0	0	4	13	0	0	9
5	0	0	0	0	5	0	0	0	2
6	0	0	0	0	6	0	0	0	0
7	0	0	0	0	7	2	0	0	0
8	0	0	0	0	8	0	0	0	0
9	0	0	0	0	9	0	0	0	0
10	1	0	0	0	10	2	0	0	0
11	0	0	1	0	11	0	0	0	0
Total	11	0	5	8		34	9	2	23
Mean	1.08	0	0.45	0.73		2.62	0.69	0.15	1.77
<i>Target failure</i>									
1					1				
2					2				
3					3				
4					4				
5					5				
6					6				
7					7				
8					8				
<i>Safety failure</i>									
1					1				
2					2				
3					3				
4					4				
5					5				
6					6				
7					7				
8					8				
9					9				
10					10				
11					11				
12					12				
13					13				

(Strauss and Corbin, 1990), codes were also identified inductively to include words or phrases found in the Assurance Frameworks. It was recognised that some of the codes, for example “access”, could apply to more than one boundary. Where this was found to be the case, the context was carefully examined before allocating to the appropriate boundary category. Certain elements of the Target Failure category, such as parts of the Standards for Better Health (DoH, 2006b), have components of patient safety within them. Where elements were explicitly safety related, such as control of infection, they were allocated to the safety failure category. Within the Financial Failure category “activity” was used to describe the financial risk of reduced activity (fewer admissions, therefore less income) due to changes in strategy by local purchasers of hospital services.

Data analysis

The content analysis involved searching for and analysing the frequency of the identified codes in each Assurance Framework document. The protocol was used to guide the inclusion, exclusion and allocation of codes to boundary themes. Where the context of the code related to an area of risk being monitored it was included in the analysis. Where the code was part of a job title, such as Director of “Finance”, or referred to a committee, such as the “Infection” Control Committee, these were excluded. Each occurrence of the code was checked against its context. Where the context related to managing a risk issue it was included in the count. The mean of all the counts was calculated for each risk boundary category.

Results

The results support the proposition that in the four hospitals the boundaries of financial and target failure feature more in their Assurance Frameworks than the other boundaries. The summary of the means for each boundary category are set out in Table II. The inclusion of risks related to staff was very limited. Patient safety did feature in general terms and more specifically in regard to infection control which has nationally set targets to be achieved. Aspects of patient safety derived from patient safety campaigns such as “medicines management” or the “deteriorating patient” did not feature. Given the known rate of “adverse events” and “harm” that occur in NHS hospitals (Sari *et al.*, 2007) it is surprising that they did not feature significantly. Whilst “quality” was used in the document, it was found to be a general term without any meaningful definition.

The results were discussed with the Director responsible for patient safety in one of the study hospitals. They confirmed the high level of focus on targets and achieving

Hospital	Finance	Target	Staff workload	Safety	Dr Foster mortality rating
A	7.0	4.38	1	2.62	High
B	1.42	7.25	0	0.69	Medium
C	1.0	19.50	0.45	0.15	Low
D	3.75	8.13	0.73	1.77	Investigated

Note: Hospitals B and C mean for “Target” is significantly higher at $p < 0.05$

Table II.
Assurance framework
content analysis – mean
for each category by NHS
hospital

the requirements of Standards for Better Health (DoH, 2006b), (Orzell, 2008).. That focus had helped the hospital to be rated as “excellent” by the Healthcare Commission for both “quality of services” and “use of resources”. Although the Assurance Framework is the key risk management document for the Trust Board according to the Department of Health (DoH, 2003), other reports and meetings were used in that Hospital to consider aspects of patient safety and staffing (Orzell, 2008).

Discussion

We can theorise that where the financial and target failure is more clearly articulated, the pressure on the operating point to keep away from those boundaries means that the system will be operating closer to the boundaries that receive more limited consideration. This can be illustrated using the descriptive model and the results from the content analyses of Hospital C’s Assurance Framework (see Figure 5). The results can do no more than illustrate where management attention is focused. However, by showing such focus it draws attention to potential blind spots or instances of normalization where the system may be more vulnerable to breaching the marginal zone.

The NHS hospitals studied do not appear to use the Assurance Framework to focus their efforts on setting a safe context for front line clinical treatment. Extant literature suggests that the pressure on the operating point at times of peaks in patient demand is a risk issue both for staff workload and patient safety (Weissman *et al.*, 2007). Human error within the context of unsafe systems is a well known phenomenon (Reason, 1990). The risks associated with high volumes of activity, their impact on staff and other

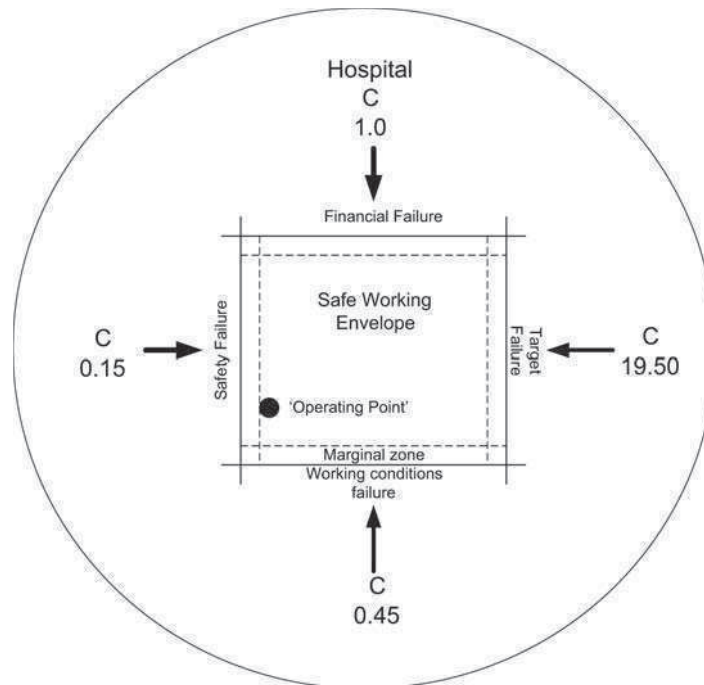


Figure 5.
A Descriptive model of a safe working envelope for NHS Hospital C

patient safety issues, does not receive the same degree of attention as finance and targets in the key risk management document for Trust Boards in this pilot study.

The management focus in the Assurance Frameworks is predominately on the corporate risks associated with financial and target failure. These are risks more associated with an impact on the organisational management and less on the clinical safety of patients. Vaughan (1996) argues that the “production of culture” can occur at the macro level and heavily influence the “normalization of deviance” at the micro level. The Department of Health, at the time of this study, had issued an “Operating Framework” which focuses on financial management and targets with almost no mention of patient safety (DoH, 2007). It is therefore perhaps not surprising that a Department of Health mandated document for risk management, the Assurance Framework used by Trusts (DoH, 2003), focused on the risk related to the priorities set by the higher authority.

Drug errors are one of the most common adverse events that occur to patients (NAO, 2005). The fact that “medicines management” or “prescribing errors” was not included in any of the Assurance Frameworks illustrates that the focus of the documents was on corporate rather than patient or clinical risk. Given the nature of hospital work it would not be unreasonable to expect that risks to patients from prescribing errors be regarded as a “principle risk” (DoH, 2003), to the achieving the objectives of the hospital.

Conclusion

The NHS hospitals studied adopt an approach to risk management at board level which focuses predominately on the corporate risk related to finance and improvement targets. Such “active” priorities have to be balanced by the “chronic” or maintenance priorities of patient and staff safety (Woods, 2006). Without this balance patient safety will remain a casualty of the emphasis on meeting other priorities and targets. The current measurement systems appear to be a product of the productivity and efficiency culture. This is inadequate in a healthcare context. Given the primary function of a hospital is first to “do no harm” to patients, there is an imbalance in the way priorities are set, measured and monitored. This pilot study provides support for the proposition, developed from the literature, that the failure boundaries for finance and targets would be more clearly set out and monitored than those associated with safety.

At an organisational level, a conceptual four boundary model of a safe working envelope can bring to mind the concept of an “operating point” and the conflicting pressures that can occur to move the operating point into the marginal zone with the potentially to breach a boundary. If the wider and dynamic pressures, such as the pressure created by peaks in demand, are not taken into account when making decisions then senior managers are in danger of developing hidden conditions that create the environment for accidents to occur (Reason, 1995). By developing the work of Cook and Rasmussen (2005) into having four, rather than three boundaries, we are taking account of the particular circumstances of a politically led NHS with a very strong target and financial control culture which can place substantial pressure on the operating point (Barber, 2008; Healthcare Commission, 2008). The idea of compensating actions constituting the marginal zone, with the potential for such actions to be measured, can be developed and tested as an indicator of the resilience concept.

For practitioners the conceptual model will assist them in recognising the conflicting pressures that they work with. Such a resilience model may also stimulate the setting of boundary measures and finding a mechanism to monitor the location and stability of the operating point at various levels within a hospital. This could potentially be done through a combination of setting boundary marking standards for all four boundaries, examining the nature and frequency of compensating actions and learning from incident reports. The Assurance Framework could be developed to provide a more balanced perspective on risks and hence design greater resilience into the system.

Limitations and future research

The pilot study using a content analysis methodology has limitations. The sample is too small to provide generalisable data and looked only at the Assurance Framework of each hospital. It was limited in the examination of other documentation. The words and phrases used, whilst developed from the literature and reading the Assurance Framework documents, cannot cover all aspects of risk management. However the exploratory research has developed the conceptual model from resilience engineering in relationship to patient safety and has highlighted fundamental issues for further research

Empirical work using an embedded case study mixed method approach is now being undertaken to build on the conceptual ideas put forward in this paper. The research is examining the setting and measurability of boundary standards and exploring the nature and measurability of compensating actions in relation to the safety boundary.

References

- Barber, M. (2008), *Instruction to Deliver: Fighting to Transform Britain's Public Services*, Methuen, London.
- Catchpole, K., Kreckler, S., Mishra, A., New, S. and McCulloch, P. (2008), "The fundamental importance of bottom-up interventions in the system of surgery", *Proceedings of Improving Patient Safety Conference, The Ergonomics Society, Cambridge, 16-18 July*, pp. 15-19.
- Cook, R.I. and Rasmussen, J. (2005), "Going solid: a model of system dynamics and consequences for patient safety", *Quality and Safety in Health Care*, Vol. 14 No. 2, pp. 130-4.
- DoH (2000), *An Organisation with a Memory. Report of an Expert Group on Learning from Adverse Events in the NHS*, Stationery Office, London.
- DoH (2003), *Building the Assurance Framework: A Practical Guide for NHS Boards*, Stationery Office, London.
- DoH (2006a), *Safety First: A Report for Patients, Clinicians and Healthcare Managers*, Stationery Office, London.
- DoH (2006b), *Standards for Better Health*, Stationery Office, London.
- DoH (2007), *Operating Framework, 2008/2009*, Stationery Office, London.
- Dr Foster (2007), *The Hospital Guide*, Dr Foster, London.
- Healthcare Commission (2008), *Learning from Investigations*, Commission for Healthcare Audit and Inspection, London.

- Institute for Healthcare Improvement (2008), *Five Million Lives Campaign*, available at: www.ihf.org/IHI/Programs/Campaign/ (accessed 10 April 2008).
- Lozeau, D., Langley, A. and Denis, J. (2002), "The corruption of managerial techniques by organisations", *Human Relations*, Vol. 55 No. 5, pp. 537-64.
- Miller, A. and Xiao, Y. (2007), "Multi-level strategies to achieve resilience for an organisation operating at capacity: a case study at a trauma centre", *Cognition, Technology & Work*, Vol. 9, pp. 51-66.
- Mohr, J.J., Batalden, P. and Barach, P. (2004), "Integrating patient safety into the clinical microsystem", *Quality and Safety in Health Care* Vol. 13, pp. ii34-ii38.
- National Audit Office (NAO) (2005), *A Safer Place for Patients: Learning to Improve Patient Safety*, Stationery Office, London.
- Olsen, S., Neale, G., Schwab, K., Psaila, B., Tejal Patel, E., Chapman, J. and Vincent, C. (2007), "Hospital staff should use more than one method to detect adverse events and potential adverse events", *Quality and Safety in Health Care*, Vol. 16, pp. 40-4.
- Orzel, M.N. (2008), personal communication.
- Proudlove, N.C., Gordon, K. and Boaden, R. (2003), "Can good bed management solve the overcrowding in accident and emergency departments?", *Emergency Medicine Journal*, Vol. 20, pp. 149-55.
- Rasmussen, J. (1997), "Risk management in a dynamic society: a modelling problem", *Safety Science*, Vol. 27 Nos 2/3, pp. 183-213.
- Reason, J. (1990), *Human Error*, CUP, Cambridge.
- Reason, J. (1995), "Safety in the operating theatre – Part 2: human error and organisational failure", *Current Anaesthesia and Critical Care*, Vol. 6, pp. 121-6.
- Reason, J. (1997), *Managing the Risks of Organizational Accidents*, Ashgate, Aldershot.
- Resar, R. (2006), "Making non-catastrophic health care processes reliable: learning to walk before running in creating high reliability organisations: Pt II", *Health Service Research*, Vol. 41 No. 4 pp. 1677-89.
- Richardson, D.B. (2006), "Increase in patient mortality at 10 days associated with emergency department overcrowding", *Medical Journal of Australia*, Vol. 184 No. 5, pp. 213-16.
- Sari, A.B-A., Sheldon, T.A. and Cracknell, A. (2007), "Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital", *Quality and Safety in Health Care*, Vol. 16, pp. 434-9.
- Sheridan, T.B. (2008), "Risk, human error, and system resilience: fundamental ideas", *Human Factors*, Vol. 50, pp. 418-26.
- Sprivilis, P.C., DaSilva, J-A., Jacobs, I.G., Frazer, A.R.L. and Jelinek, G.A. (2006), "The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments", *Medical Journal of Australia*, Vol. 184 No. 5, pp. 208-12.
- Strauss, A. and Corbin, J. (1990), *Basics of Qualitative Research*, Sage, Newbury Park, CA.
- Vaughan, D. (1996), *The Challenger Launch Decision*, Chicago University Press, Chicago, IL.
- Vincent, C., Neale, G. and Woloshynowych, M. (2001), "Adverse events in British hospitals: preliminary retrospective record review", *British Medical Journal*, Vol. 322, pp. 517-19.
- Vincent, C., Taylor Adams, S. and Stanhope, N. (1998), "Framework for analysing risk and safety in clinical medicine", *British Medical Journal*, Vol. 316, pp. 1154-7.
- Waring, J.J. (2005), "Beyond blame: cultural barriers to medical incident reporting", *Social Science & Medicine*, Vol. 60, pp. 1927-35.

- Waring, J., Harrison, S. and McDonald, R. (2007), "A culture of safety or coping? Ritualistic behaviours in the operating theatre", *Journal of Health Services Research and Policy*, Vol. 12 (Suppl. 1), pp. 3-9.
- Weissmann, J., Weissmann, J., Rothschild, J.M., Bendavid, E., Sprivulis, P., Cook, E.F., Evans, R.S., Kaganova, Y., Bender, M., David-Kasdan, J., Haug, P., Lloyd, J., Selbovits, L.G., Muff, H.J. and Bates, D.W. (2007), "Hospital workload and adverse events", *Medical Care*, Vol. 45 No. 5, pp. 448-55.
- Woods, D.D. (2006), "Essential characteristics of resilience", in Hollnagal, E., Woods, D. and Leveson, N. (Eds), *Resilience Engineering: Concepts and Precepts*, Ashgate, Aldershot.
- Woods, D.D. and Cook, R.I. (2002), "Nine steps to move forward from error", *Cognition, Technology & Work*, Vol. 4, pp. 137-44.
- Woods, D.D. and Cook, R.I. (2006), "Incidents – markers of resilience or brittleness?", in Hollnagal, E., Woods, D. and Leveson, N. (Eds), *Resilience Engineering: Concepts and Precepts*, Ashgate, Aldershot.
- Woods, D.D. and Hollnagal, E. (2006), "Resilience engineering concepts", in Hollnagal, E., Woods, D. and Leveson, N. (Eds), *Resilience Engineering: Concepts and Precepts*, Ashgate, Aldershot.

Further reading

Royal Devon and Exeter NHS Foundation Trust (2007), *Risk Management Strategy*, Royal Devon and Exeter NHS Foundation Trust, Exeter.

About the authors

Mike Dermot Williams is currently NHS Senior Research Fellow at University of Exeter Business School and a member of the core team for the National Patient Safety Campaign in the UK. Prior appointments have been as Chief Executive and Director of NHS Teaching and District Hospital Trusts in the UK. Mike Dermot Williams is the corresponding author and can be contacted at: m.d.a.williams@ex.ac.uk

Andi Smart is a Research Fellow at the University of Plymouth. From 1998-1999 he was Research Fellow, Department of Business and Management, School of Business and Economics, University of Exeter. From 1999-2003 he was Lecturer in Management, Department of Business and Management, School of Business and Economics, University of Exeter. From 2003-present he has been Senior Lecturer in Operations and Process Management, Exeter Centre for Research in Strategic Processes and Operations (XSPO), School of Business and Economics, University of Exeter.

Publication 2

Williams, M. D. & Smart, P. A. (2009) System resilience and patient safety during a bed crisis in an NHS hospital in England. *16th Annual EurOMA Conference, Goteborg.*

System resilience and patient safety during a bed crisis in an NHS hospital in England

Mike D Williams

Andi Smart

m.d.a.williams@exeter.ac.uk

University of Exeter Business School

Rennes Drive

Exeter

EX4 4PU

Abstract

A resilience based conceptual model of a four boundary safe working envelop was developed and used in a case study in a NHS hospital to develop theoretical explanations about the actions and decisions of staff in terms of how they deal with competing requirements. A mixed method of data collection was used during a bed crisis. Evidence of a process of normalization relating to capacity problems was found; a focus on what was easily measured in terms of finance, waiting time targets, and staffing predominantly drove operational behaviour. Patient safety may be compromised by the process of normalisation through the theoretical concept of ‘drift to danger’.

Keywords: resilience, normalisation, patient safety

Introduction

Hospital managers and clinicians face many competing demands. These include the need to meet patient expectations in terms of effective, safe and timely treatment whilst meeting the organisational expectations, including remaining within budget. The publically funded healthcare system in the UK, through the government, sets hospitals a number of standards and waiting time targets that hospitals have to meet. A conceptual model, derived from resilience engineering, was developed to help explain how staff actions keep hospital patients safe during periods of high levels of patient demand and competing pressures (Williams, 2008). This model was applied within a case study of a UK hospital to develop a systems resilience theory to explain the actions and decisions of managers and clinicians in relation to patient safety during a period of bed capacity constraint and high patient demand. Patient safety in hospitals throughout the world is a major issue with adverse events occurring to over 9% admitted patients (De Vries et al, 2008). Developing theory that will assist in reducing harm to patients during busy periods in hospitals has potential practical impact.

Literature

Resilience engineering is a developing field within the literature (Hollnagel et al, 2006). It seeks to take a systems and proactive perspective on safety. The definition of ‘system resilience’ in this paper relating to healthcare is the ‘work system’s ability to buffer, adapt to, absorb and prevent adverse patient outcomes in the face of disruption’ (Woods, 2006).

The resilience literature includes the idea of a ‘safe working envelop’ within which a work system seeks to remain to avoid failure. Rasmussen (1997) describes three interacting boundaries to the safe working envelope; the boundary of economic failure, the boundary of unacceptable workload and the safety boundary of unacceptable performance. The model includes a marginal boundary or zone inside those boundaries which if breached would create the conditions for failure.

Cook and Rasmussen (2005), use Rasmussen’s safe working envelope model to examine safety problems for hospitals that are overfull. They suggest that it is normal for healthcare systems to work at the limit of their capacity and for Hirschhorn’s ‘law of stretched systems’ to apply: ‘every system always operates at its capacity. As soon as there is some improvement, some new technology, we stretch it...’. When that capacity is reached, failure or restructuring has to occur (Woods and Wreathrall, 2008). When systems are under resource and or performance pressures the benefits of change are taken in increased productivity or efficiency. Studies have shown an association, but not proven causation, between hospitals that have high occupancy rates and an increase in mortality and adverse events (Richardson, 2006; Sprivulis, et al, 2006; Weissman et al, 2007).

Building on the work by Cook and Rasmussen (2005) and Miller and Xiao (2007), we developed a conceptual model to explain system resilience of NHS hospital systems in the wider political and administrative context by proposing four interacting boundaries of a ‘safe working envelope’ (Williams, 2008). These are the boundaries of financial failure, target failure, unacceptable working conditions, and failure of safety (Figure 1). This model seeks to explain the dynamic pressures that apply to the concept of a safe working envelope for NHS hospitals.

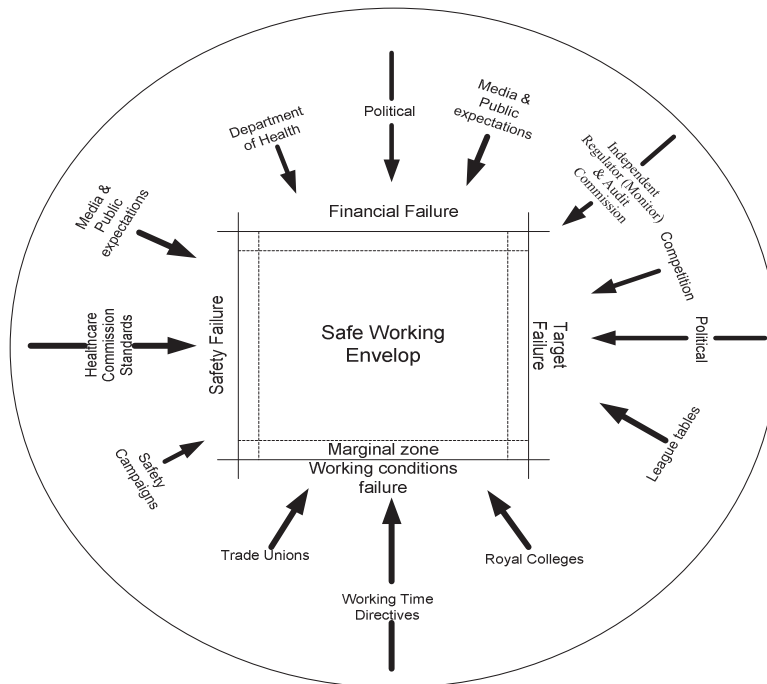


Figure 1: A conceptual model of a ‘safe working envelope’ for an NHS hospital

Vaughan (1996) shows that normalisation of deviance is key issue for safety. Her study of the Challenger Space Shuttle disaster illustrated that the system within which people work can produce a culture where, through small incremental steps, new situations are seen as 'normal'. This is described as the 'native view'. Outsiders looking at the situation are more likely to see the situation as deviant, not acceptable and therefore potentially dangerous. Waring et al (2007) point to types of behaviour by health care workers that illustrate 'taken for granted assumptions about clinical risk...' Such behaviours 'normalize risk' and as such mean that risk of harm to patients is not addressed. Waring (2005) suggests that medical staff regard error as 'inevitable' which can lead to errors being seen as 'normal'. These errors may not be reported as incidents and therefore do not provide the basis for improvements in safety. Normalization in the context of the safe working envelop can be regarded as a 'drift to danger' of the operating point. Equally, normalization can be conceived of as shifting the safety failure boundary to a new but more dangerous position.

Method

A case study approach was used in an NHS hospital during an outbreak of a sickness virus. The case study approach was chosen because it is suitable to investigate phenomenon within their context Yin (2003) and where there is complexity in the subject matter (Stuart et al, 2002). It is also used where the boundaries between the phenomenon are unclear, where multiple sources of evidence are needed to converge, and where there is a theoretical background to guide the data collection and analysis (Yin, 2003). Case studies do suffer from the weakness of being context specific and therefore transferability of findings to other situations is problematic. However, when case studies are used to draw out key points of an explanation they can be a powerful contribution (Nightingale et al, 2003).

The case was chosen as an example of a high performing teaching hospital with a stable leadership team where the externally validated track record would suggest an ability to manage both the external and internal competing pressures. A period of intense pressure on bed capacity (a sickness virus that closed wards and increased staff absence) was studied as a means of magnifying the competing pressures, the staff actions in response to those pressures and the implications for patient safety. Full NHS ethical approval was gained.

Within the case study the unit of analysis is the whole hospital as a system. There is an embedded approach (Yin, 2003), using a stratified purposeful approach (Miles and Huberman, 1994) to examine three within case units of analysis to build a picture of the wider hospital. The reason for choosing an embedded approach is to overcome the weakness of a holistic case study design. Yin (2003) argues that a holistic approach may mean the study is conducted at an abstract level as the researcher may not get into the operational detail. One of the reasons behind selecting three embedded areas is to cover both the formal and informal actions that occur, to understand the macro context and then the managers and individual team members (micro) response. The three units were:

Organisational:	Trust Board and hospital wide operational processes
Sub unit:	Division of Medicine
Team:	Ward / Consultant Team / Department members

A mixed method approach using the simultaneous collection of qualitative and quantitative data was used (Creswell, 2003; Bergman, 2008). The underlying pragmatic philosophy is that to obtain a better understanding of the phenomenon, a range of epistemological perspective were used. Multiple sources of data collection were utilized.

This assists in achieving triangulation in data collection, the interpretation and subsequent theory development. A concurrent nested approach to gathering data through quantitative and qualitative means was employed to allow different questions to be addressed during the study. Such an approach allowed different perspectives to be obtained about the hospital being studied and by bringing the data together during the analysis phase. This helped to increase the validity of the findings (Creswell, 2003). The collection of quantitative data was nested within and informed the wider qualitative study (Creswell and Plano Clark, 2007).

Data was collected from a number of sources. Initially there was analysis of both internal and external context setting documents to provide the contextual overview. Hospital administrative data provided the descriptive statistics relating to demand and capacity before and during the sickness virus. During and after the period of high pressure semi-structured interviews of eight doctors, nine nurses, ten managers were conducted, recorded and transcribed. Non-participative observation of staff actions and meetings were undertaken over a four month period. A grounded theory approach was used in the analysis to develop categories and concepts from the different data sources using NVivo (Strauss and Corbin, 1990). Thematic analysis of the concepts was used to interpret the data to identify phenomena about the competing priorities and the failure boundaries (Miles and Huberman, 1994).

Case study hospital

The hospital studied is a 760 bed NHS teaching hospital serving a largely rural population in England. The nearest acute hospital is over twenty miles away. The main targets that impact on hospital bed capacity are (DoH 2007):

- A maximum of waiting time of one month from diagnosis to treatment for all cancers.
- Waiting time in Emergency Department (ED)– maximum of 4 hours for 98% of patients
- Waiting time from Referral To Treatment (RTT) – maximum of 18 weeks with a push to reduce to 15 then 13 weeks (This target relates to non-emergency patients)

Patient routes into the hospital

Where cancer is suspected then the referring General Practitioner (GP) uses a rapid access route to ensure that the patient has an outpatient appointment within 14 days and then, where necessary, treatment commences within 31 days.

Emergency patients enter the hospital through the ED, or if they have been seen by a GP prior to attending hospital, they will be sent directly to the surgical assessment unit (SAU) or emergency medical unit (EMU). Ninety eight percent of those patients arriving at the ED must be treated and admitted or discharged from ED within four hours. There are no targets relating to patients sent directly to the SAU or EMU.

Non-emergency and non cancer patients are referred by their GP to a consultant in the hospital. The usual pathway is for a consultation in an outpatient clinic either before or after some diagnostic tests. Once a decision on treatment is made, the patient may be admitted as a day case or inpatient for treatment. Ninety percent of such patients being admitted have to have their treatment started within the RTT target.

The three target areas can create competing dynamics within a hospital service. For the ED to succeed on admitting patients within 4 hours depends on other parts of the hospital having the capacity at the time required. However, at the same time the SAU/EMU receives patients directly from GPs. Once these patients have been assessed, they need to be accommodated in suitable specialty wards (Proudlove *et al*, 2003). Those patients on the RTT pathway also have to be admitted within the target of 18 weeks and also consume bed capacity.

The high pressure event

The case study focused on before, during and after a specific incident in the hospital. An airborne sickness virus in one ward that then spread to other patients and staff. Within a short time the virus was present in eight wards (around 30% of total bed capacity). The Control of Infection Team procedures for such an outbreak means the affected wards were closed to new admissions. The reduction in bed capacity continued for the following two weeks with ward-based capacity being reintroduced following sterilization and confirmation of the elimination of the virus.

Findings

Identifying the four boundaries of the safe working envelop

Analysis of the performance management documentation and interviews showed that both the financial and target boundaries were easily measured, monitored and strictly performance managed. The financial performance was reviewed at least monthly in substantial detail. The four hour target was monitored continuously at busy times and other targets at least weekly. Staff numbers were linked to budgets and monitored monthly for turnover and sickness rate. Schedules of expected numbers of doctors and nurses per ward per shift were in place and staffing was monitored informally each day. The measurement of patient infections was undertaken daily and regarded as the primary patient safety issue. 'Patient falls' and incident reports were formally monitored retrospectively each month. Within certain areas senior nurses reviewed incidents of patient falls within a week. Apart from infections and number of serious incidents no other patient safety measures were reported to the Hospital Board on a routine basis. Financial and target performance were reported to the Board in detail each month. Staff turnover and sickness rates were reported every quarter.

Pressure on the 'operating point'

Theoretically 'operating point' is the location of the system within the safe working envelop. Dynamic pressures influence the location of the operating point and can push the system towards a boundary of failure. The case study hospital was in a strong financial position with a projected large surplus for the year. The interviews confirmed that staff did not feel pressure to make decisions that might compromise targets, staffing or patient safety due to budget considerations.

The combination of the four hour ED, Cancer and RTT targets did create substantial pressure. Any NHS hospital has to be able to manage peaks in emergency demand whilst at the same time maintain the pattern of non-emergency (elective) admissions from both the cancer and RTT pathways. Bagust *et al* (1999) showed that for a hospital to be in a position to manage peaks in demand the bed capacity should not exceed 85% occupied.

The hospital managers at the outbreak of the sickness virus had taken the decision that it was not feasible to divert emergency patients. Therefore, the remaining potential action available was to cancel non-emergency admissions to help relieve pressure. From the interviews there appeared to be a strong view that non-emergency patients are equally deserving of admission.

“The non-emergency patient has a problem that needs treatment. ...From the patient’s perspective a bed crisis is not their problem – they just want to be admitted.” (Nurse)

As well as the patient perspective from observations and interviews it was clear that to cancel an elective patient on the RTT pathway created a capacity problem in trying to rebook the patient within the target period. Also there was a government target of the percentage (<1%) of patients that could be cancelled on the day of admission for non-medical reasons. The competing dynamics led to a process of tight coupling where operational problems in one part of the hospital impacted adversely on other areas (Cook and Rasmussen, 2005). This generated considerable pressure on staff to find innovative ways to maintain the flow of patients through the system by seeking to accelerate patient discharges.

Data analysis showed there was no significant change in the level of elective admissions during the bed capacity crisis (Figure 2) compared to the same period the previous year. Without a change the level of non-emergency admissions, other actions relating to capacity, particularly maintaining the flow of patients, became the focus for attention.

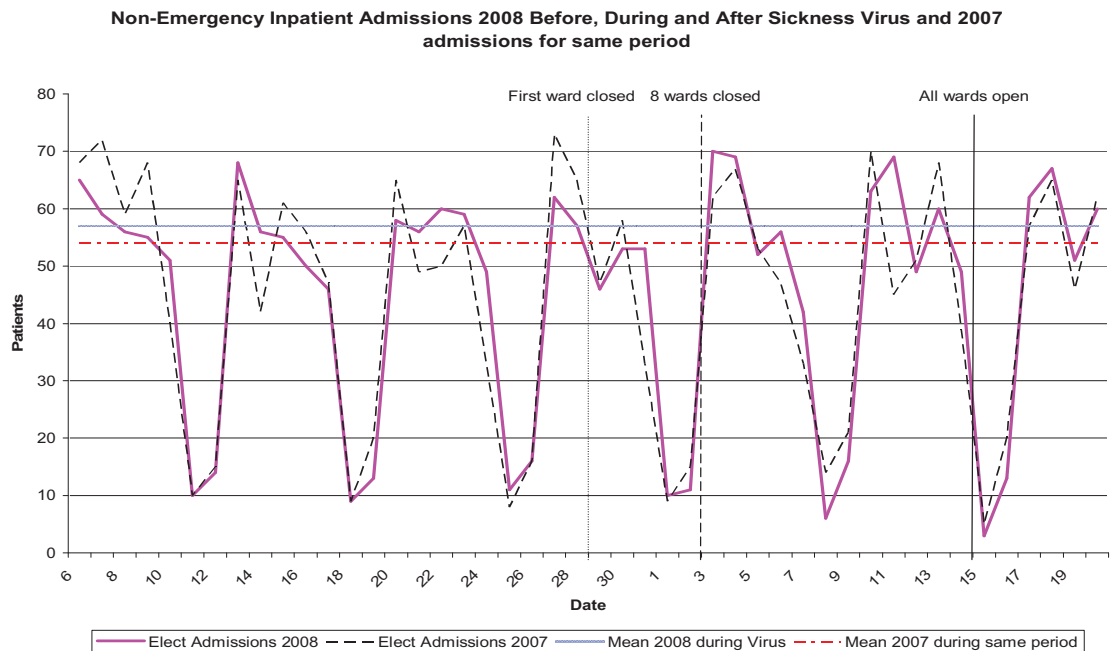


Fig 2 Elective Admissions 2008 Before, During and After Sickness Virus and 2007 over same period

A number of actions were taken to generate new staffed bed capacity (overflow capacity). This included opening an area normally used as a preadmission area (normally only open during the working day) in orthopedics. This small 13 bed ward is situated at the other end of the hospital from the medical wards. Patients already in medical beds, identified as being well enough, were then moved to this area. A further area of 6 beds was brought into use in an area close to the ED for medical patients. Areas within wards that were normally used as day case beds were brought into service as inpatient beds.

Staffing the additional areas combined with the short-term sickness of doctors and nurses due to the virus pushed the operating point towards the staffing boundary and an inadequate level of staff in ward areas. Concerns were expressed about staff having to move to cover sickness and open additional beds. Staff were unfamiliar with certain types of patients and ward areas which lacked key equipment.

Infection control measures were applied rigorously. Wards with the sickness virus remained closed until all the patients and staff were symptom free for forty eight hours and then the areas thoroughly cleaned. Emergency patients admitted with an infection were placed in a side room. These actions maintained the patient safety boundary in respect of infection but then created pressure on the remaining bed and staff capacity. That pressure on the remaining bed capacity meant that medical patients were transferred onto surgical wards (medical outliers). By moving patients into surgical beds it provided the capacity within the medical wards to accept transfers from EMU. In turn then EMU can accept admissions from the ED. With the 4 hour target there has to be a constant supply of beds being found in the hospital. During the sickness virus period both the EMU and sometimes the SAU would not have any empty beds. Emergency patients requiring admission after seeing their GP, were then diverted to the ED. This created a reinforcing loop (Sterman, 2000) by creating more patients to be processed within the four hour target which meant the need to create more empty beds in the hospital within a short time frame. As the majority of emergency admissions were medical patients, current medical patients were moved to surgical wards (medical outliers) in increasing numbers to create the bed capacity in the 'right place'.

“Sometimes you have two or three bed moves to create a bed on EMU to get an ED patient in. ...The medical patients that move are not always suitable.” (Nurse)

As suggested by Wheeler (2006), a moderate but sustained effect on the number of medical outliers was observed during the virus period (Figure 3).

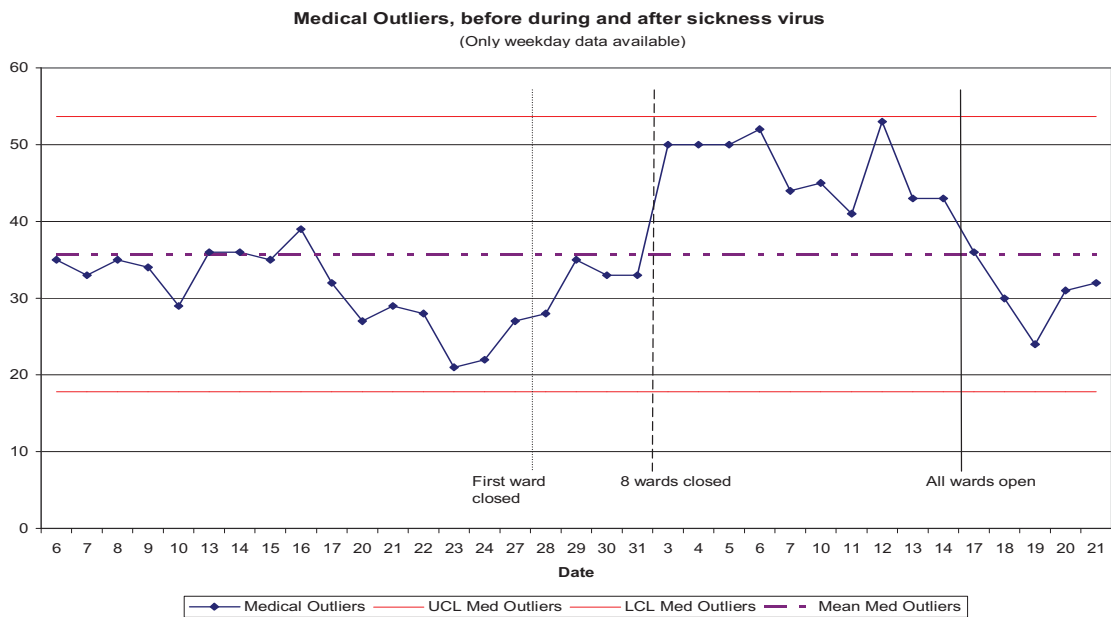


Fig 3 Statistical Process Control Chart - Medical outliers before, during and after sickness virus

Interviews and observation confirmed that bed capacity and flow problems were common and regarded as ‘normal’. There was evidence of admitting and preparing non-emergency patients for procedures in corridors and waiting rooms. NHS hospitals count the bed occupancy at midnight. In 2008 the midnight occupancy for all wards in the case study hospital was 89%. The medical wards occupancy ranged from 90-98%. The midnight census underestimates the bed occupancy during peak period of activity during the day. Occupancy for many wards during the day was above 100%. Staff had learnt to accept the pressure on the system and had found innovative ways to manage the situation.

Discussion

From the literature on safety theory and stretched systems (Cook and Rasmussen, 2005; Miller and Xiao 2007; Rasmussen, 1997; Reason, 1997; Sheridan, 2008; Woods and Cook, 2002, Woods, 2006), we can conceptualize that the operating point would be close to the safety boundary for a hospital experiencing capacity and production pressures. In the case study only three out twenty seven staff interviewed believed the operating point was within or near the marginal zone. Most regarded the hospital as operating in a way that prioritized the safety of patients. Previous (albeit limited) studies suggest that hospitals that have high occupancy rates, experience an increase in mortality and adverse events (Richardson, 2006; Sprivilis, *et al*, 2006; Weissman *et al*, 2007). It is therefore surprising that more staff did not regard the hospital as being unsafe during this period. It can be argued that the hospital system experienced a period of ‘drift to danger’ (Rasmussen, 1997).

There was evidence that the buffer capacity and flexibility required for a resilient response was present only on the financial failure boundary. The admission process in terms of beds and staff became brittle (no flexible capacity) and tightly coupled (Woods, 2006; Cook and Rasmussen, 2005).

Role of culture

Vaughan (1996) describes how the decision to launch the ill fated shuttle, Challenger, was as much about culture as it was about technical systems going wrong. Weick and Sutcliffe (2003) point to the importance of a safety culture, defining culture as ‘what we expect around here’. There was considerable evidence from the observation and interview data of staff normalizing to the pattern of actions that maintained the level of admissions. They appeared to make sense of the situation by placing importance on not canceling patients who needed non-emergency admissions. We can theorise that the reality was more likely the system pressure on staff to achieve multiple targets. As Vaughan (1996) suggests: ‘We reconstruct history every day, not to fool others but to fool ourselves, because it is integral to the process of going on.’

In the NHS waiting times and measurable infection control targets have been given a high priority backed by a strong performance management ethos. This contributes to a culture of finding all possible ways to meet the targets. The wider NHS culture also allowed a process of normalizing by staff to levels of risk previous thought to be unacceptable. Safety being a dynamic non-event (Reason, 1997), and not as clearly defined or measured as finance and waiting time targets has received far less management attention. The second aspect is the culture of production. Staff in the hospital had normalized to being extremely busy to compensate for the constant pressure of competing demands and the need to find beds for admissions. The focus on maintaining flow of patients through the and out of the hospital has become hard wired into the system in terms of staff attitudes, processes and structures.

Conclusion

The conceptual model of a safe working envelop for an NHS hospital has provided the framework through which to explore the compensating actions staff take in the face of competing pressures. The research has led to the articulation that the system pressures and failure boundaries influence the attitudes of staff and consequently the culture which drives the actions. This conclusion may relate to what Weick and Sutcliffe (2003) describe a process of cultural entrapment – ‘...the process by which people get locked into lines of action, subsequently justify those lines of action, and search for confirmation that they are doing what they should be doing.’

The need for the safe working envelop model to include the wider system is needed to conceptualize the impact of externally generated pressures both on the production of culture and the culture of production. To extend this approach, to explain the impact of conflicting pressures on patient safety, further research incorporating additional cases is needed. A single case study has limitations. The situation of an NHS hospital facing the pressures described is particular to England. With those limitations in mind the development of the conceptual model to a particular system wide context, does facilitate the articulation of compensating actions as a response to competing pressures.

From a practitioner perspective the model helps illustrate the dynamic nature of the competing pressures and the need to set clearer boundary measures in the area of staff workload and patient safety. Policy makers, performance managers and regulators may be able to appreciate better the impact their decisions have on the actions of front line staff in hospitals and the potential impact this, in turn, has on patient safety.

References

- Bagust, A., Place, M., and Posnett J.W. (1999), Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *BMJ*, Vol. 319, pp.155-8.
- Bergman M.M. Ed (2008), *Advances in Mixed Methods Research*, Sage, London.
- Cook, RI, Rasmussen J. (2005), "Going solid": a model of system dynamics and consequences for patient safety. *Qual Saf Health Care*, 14(2): 130-134.
- Creswell J.W. (2003), *Research Design* (2nd ed) Sage, London
- Creswell J.W, and Plano Clark V.L. (2007), *Designing and Conducting Mixed Methods Research*, Sage, London.
- Department of Health. (2007), *Operating Framework 2008-09 for the NHS in England* available at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081094 (accessed 24 November 2008)
- De Vries E., Ramrattan M., Smorenburg S., Gouma D., and Boermeester M. (2008), The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care*, 17: 216-223.
- Hollnagal E, Woods D, Leveson N. (eds) (2006), *Resilience Engineering: concepts and precepts*, Aldershot, Ashgate.
- Miles M.B, and Huberman A.M. (1994), *Qualitative Data Analysis* (2nd ed), Sage, Thousand Oaks, CA.
- Miller A, Xiao Y. (2007), Multi-level strategies to achieve resilience for an organisation operating at capacity: a case study at a trauma centre. *Cogn Tech Work*, 9: 51-66
- Nightingale, P., Brady, T., Davies A. and Hall, J.(2003), Capacity utilization revisited: software, control and the growth of large technical systems. *Industrial and Corporate Change*, 12(3):477-517.
- Proudlove NC, Gordon K, Boaden R. (2003), Can good bed management solve the overcrowding in accident and emergency departments? *Emerg Medicine Journal*, 20:149-155.
- Rasmussen J. (1997), Risk Management in a dynamic society: a modelling problem, *Saf Sci*, 27(2/3) 183-213
- Reason J. (1997), *Managing the Risks of Organizational Accidents*, Ashgate, Aldershot.
- Richardson, D. B. (2006), Increase in patient mortality at 10 days associated with emergency department overcrowding. *Medical Journal of Australia*, 184(5):213-216
- Sheridan T.B. (2008), Risk, Human Error, and System Resilience: Fundamental Ideas, *Human Factors*, 50, 418-426.
- Sprivulis, P. C., Da Silva J-A, Jacobs IG. et al. (2006), The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments. *Medical Journal of Australia*, 184(5): 208-212
- Sterman J.D. (2000), *Business Dynamics*, Irwin McGraw-Hill, Boston.
- Strauss A, and Corbin J.M. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Newbury Park, California.
- Stuart, I., Mccutcheon, D., Handfield, R., Mclachlin, R. and Samson, D. (2002), Effective case research in operations management - a process perspective. *Journal of Operations Management*, 20 (5): 419-33.
- Vaughan D. (1996), *The Challenger Launch Decision*, Chicago University Press, Chicago.
- Waring, J.J. (2005), Beyond blame: cultural barriers to medical incident reporting. *Social Science & Medicine*, 60, 1927-1935.
- Waring J, Harrison S, McDonald R. (2007), A culture of safety or coping? Ritualistic behaviours in the operating theatre. *Journal of Health Services Research and Policy*, 12 (Suppl 1):3-9.
- Weick K, Sutcliffe K. (2003), Hospitals as Cultures of Entrapment: A Re-Analysis of the Bristol Royal Infirmary. *California Management Review*, 48 (2):73-84.
- Weissmann J, et al (2007), Hospital Workload and Adverse Events, *Medical Care*, 45(5) 448-455
- Wheeler D.J. (2003), *Making Sense of Data*, SPC Press, Knoxville, Tennessee.
- Woods DD (2006), Essential Characteristics of Resilience. In: Hollnagal E, Woods D, Leveson N. (eds) *Resilience Engineering: concepts and precepts*. Aldershot, Ashgate pp.21-34.
- Woods D.D, Cook RI. (2002), Nine Steps to Move Forward from Error, *Cognition, Tech & Work*, 4:137-144.
- Woods D.D, and Wreathall J. (2008), Stress-strain Plots as a Basis for Assessing System Resilience. In Hollnagal E, Nemeth C, Dekker S. (eds) *Remaining Sensitive to the Possibility of Failure*. Aldershot, Ashgate, pp.143-158.
- Williams M.D. (2008), Designing System Resilience in NHS Hospitals. *Proceedings of the 15th EUROMA Conference*, EUROMA, Groningen.
- Yin R.K. (2003), *Case Study Research, Design and Methods*, Sage, London.

Publication 3

Williams, M.D, George, B. & Wade, R. (2009) Rebellion against the 'normal' to improve safety for patients. *16th Annual EurOMA Conference*, Goteborg.

Rebellion against the ‘normal’ to improve safety for patients

*Mike D Williams
Bernadette George*
Ros Wade**

*m.d.a.williams@exeter.ac.uk
University of Exeter Business School
Rennes Drive
Exeter
EX4 4PU*

**Royal Devon & Exeter NHS FT
Wonford Hospital
Barrack Road
Exeter
EX2 5DW*

Abstract

Patients who fall in hospital can suffer significant harm. This case study in an NHS hospital using interviews, observation and content analysis looked at how nurse leaders ‘rebelled’ against the norm of accepting patient falls as inevitable. A checklist approach combined with the intentionally checking high risk patients each hour was introduced. Initial results were good but not sustained. Obstacles to improvement included many competing demands, high levels of incremental normalisation supporting the theoretical position of practical drift away from set operational policies with patient safety implications.

Key words: patient falls, intentional rounding, checklist

Introduction

There are a significant number of patients who suffer adverse events in healthcare systems (De Vries et al, 2008). An ‘adverse event’ is defined as ‘any unintended event caused at least partly by healthcare and which resulted in harm’ (Sari et al, 2007). One of the most common causes of harm to patients in hospital are when they fall. A fall is defined as ‘all situations in which a patient suddenly and involuntarily came to rest upon the ground or surface lower than their original station’ (Oliver et al, 1997). In England, patient falls are more routinely reported as incidents than medication errors or adverse drug reactions (NAO, 2005). Historically staff in hospitals regarded patient falls as one of the complications that arise in hospitals and that there was little preventative action that could be taken. Therefore, a relatively high number of patient falls in hospital was regarded as ‘normal’. However, in recent years there has been a change of attitude and approach. This paper reports the outcome of case-based research in an NHS hospital, which has been undertaken to articulate the factors instigating this change of attitude. The study also identifies the specific interventions and actions undertaken by staff at the hospital to prevent patient falls.

Literature

Patient falls in hospital are common. The reported rate ranges from 3 to 14 per 1000 bed days, and are associated with a range of injuries and an increased length of stay in hospital (Healey et al, 2008, Oliver et al, 2007) In England from September 2005 to August 2006, some 200,000 falls were reported to the National Patient Safety Agency. Analysis of the

reported incidents of falls shows that 65% of patients resulted in no harm, 31% in 'low harm', 3.6% in 'moderate harm' and 0.6% in severe harm with 26 reported deaths (Healey et al, 2008). However, it is well recognised that there is serious under reporting of incidents in hospital (Olsen et al, 2007).

A systematic review and meta-analysis of studies to prevent falls and fractures suggested that there 'is reasonable evidence that using a multifaceted intervention for hospital inpatients may have a modest effect on falls but not on fractures' (Oliver et al, 2007). Interventions included risk assessment, medication review, care planning, and education. The literature supports the view that patient falls is a complex issue and that there appears to be no simple solution (NICE, 2004).

Vaughan (1996) shows that normalisation of deviance is key issue for safety. Her study of the Challenger Space Shuttle disaster illustrated that the system within which people work can produce a culture where, through small incremental steps, new situations are seen as 'normal'. This is described as the 'native view'. Outsiders looking at the situation are more likely to see the situation as deviant, not acceptable and therefore potentially dangerous. Waring et al (2007) point to types of behaviour by health care workers that illustrate 'taken for granted assumptions about clinical risk...' Such behaviours 'normalize risk' and as such mean that risk of harm to patients is not addressed. Waring (2005) suggests that medical staff regard error as 'inevitable' which can lead to errors being seen as 'normal'. These errors may not be reported as incidents and therefore do not provide the basis for improvements in safety. Grenier-Sennelier et al (2002) suggests that in hospitals, staff can regard patient falls as an inevitable part of the rehabilitation process. Such a view may then influence the reporting of such incidents (Healthcare Commission, 2009).

Merton (1968) provides a framework of how social structures exert pressure on individuals to conform to patterns of cultural goals and institutional norms. 'Rebellion' is one of the five modes of individual adaptation to goals and norms proposed by Merton (1968). The rebellion by key leaders to goals and norms is used as the theoretical explanation of the change of attitude creating a reversal of the normalisation of deviance.

There is a developing science and methodology of improvement in healthcare. It is largely based around the idea of small change which is then tested and refined before being spread to generate larger scale change in practice. The bedrock method is the cycle of plan, do, study, act (PDSA), (Langley et al, 1996).

Research questions

What were the drivers that caused the staff to change their view of 'normal' in relation to patient falls? What and how did the staff change their behaviour and practice to improve the safety of patients?

Method

A case study approach was used as being the most appropriate to gather the contextual detail needed (Yin, 2003) and where there is complexity in the subject matter (Stuart et al, 2002). A criticism of case studies is that of being context specific and therefore transferability of findings to other situations can be problematic. However, when 'they are used to contextualize key points of an explanation, and the explanation, and not the cases, carries the argument, case studies can be very powerful (Nightingale et al, 2003).

A mixed method approach using the simultaneous collection of qualitative and quantitative data was used (Creswell, 2003; Bergman, 2008). The underlying philosophy is that to obtain a better understanding of the phenomenon a range of epistemological perspectives were used. Multiple sources of data collection were utilized. This assists in achieving triangulation in data collection, the interpretation and subsequent theory development.

An 760 bed NHS hospital was chosen as the case study where an improvement methodology was being used and significant change was beginning to occur in the level of falls prevention. The case studied was five medical wards where a new approach to falls was being implemented set within the context of the wider hospital. Those wards specialised in care of the elderly (48 beds), stroke (30 beds) and neurology (24 beds).

The hospital reported 1631 in-patient ward falls in 2008. The five wards had 568 (34%) of that total. On admission patients were assessed for the risk of falling. A scoring system of over 30 indicates that the patient is at very high risk of falling. The hospital Falls Management Policy in place at the beginning of 2008, required that certain actions must be taken by staff for any patient with a risk assessment score greater than 13.

Twenty seven staff (doctors, nurses and managers) were interviewed as part of a wider study using an interview protocol and then six further semi structured interviews were conducted with nurses working within the five wards focusing on the issue of patient falls. Meetings about patient falls were observed, papers and reports analysed and data on falls plotted using statistical process control charts (Wheeler, 2003). The interviews were transcribed and then coded using NVivo 8 to develop themes which were triangulated with the observational data and content analysis of meeting papers and reports.

Findings

The analysis showed that during the first part of 2008 there were different histories within the wards of staff attitudes to patient falls. The histories were related to the leadership of the ward areas. On two of the wards (A & B) the attitude was described as falls “just being one of those things that happens in hospital”. Risk assessments often were not done and when they were, little if any action was then taken. There had been a gap in the senior nursing leadership on those wards and for other nurses the situation was described as rather “chaotic”. The other three wards (C – E) had more stable senior nursing leadership at ward level. The attitude to falls was that they were “not what you want to happen but an inevitable part of getting patients mobile”. Risk assessments were undertaken and actions put in place especially when mobilising patients. There was little if any awareness of the national initiatives to reduce harm from patient falls (DoH, 2001; NICE, 2004).

At the end of 2007 and within the first few months of 2008 on wards A&B there were two falls incidents where both patients subsequently died. Root cause analysis (RCA) of the incidents were carried out. In one case a confused patient climbed over the bed rails (raised in this case inappropriately to help prevent the patient getting out of bed) on two occasions the same night and had fallen. That patient subsequently died from a head injury sustained from the second fall. The other incident concerned a confused patient who had fallen some twenty times over a number of weeks on the ward without any significant action being taken. That patient sustained a fractured hip and subsequently died. Both RCAs showed systemic failure within the wards. The hospital Falls Management Policy did not appear to

be well known by staff and therefore poorly implemented and supervised. This situation was set within the context of the hospital achieving the highest rating for its services from the Healthcare Commission and with a Hospital Standardised Mortality Rate (HSMR) below the expected rate.

Following the RCAs the Lead Nurse with overall responsibility for all the medical wards in the hospital felt ‘ashamed that these patients had died’ and then led the ‘rebellion’ against regarding patient falls as a ‘normal’ part of a hospital stay. Her primary lever was to engage with the emotions of staff by presenting the patient stories. Most staff were not fully aware of the potential level of harm that can result from a patient fall. Staff were also shocked and some visibly upset that these incidents could have occurred. There were also many competing demands on staff time and preventing falls was regarded as just one among a number of priorities.

QUALITY: Patient Safety–Falls

For patients with a falls risk score of >30, please enter either ‘A’ = achieved or ‘V’ = variance in columns.
Record reason for variance and action taken overleaf.

THIS PATIENT REQUIRES OBSERVATION EVERY MINUTES

DATE:	TIME					
1.CONTINENCE Do you need to go to the toilet?						
2.PAIN Do you have any pain?						
3.ORIENTATION – fully alert=FA; mildly confused/disorientated=MC; severe confusion/disorientation=SC; asleep=A						
4.POSITION / COMFORT Are you comfortable?						
5.DRINK / MOUTHCARE Would you like a drink?						
6.CALL BELL WITHIN REACH If you need me, press this button						
7.BED RAILS DOWN						
8.BED TO FLOOR						
9.IS THERE ANYTHING ELSE I CAN DO? Because I’ve got the time						
INITIALS						

Figure 1 Intentional Rounding (IR) checklist

As well as engaging with the emotions of staff to shock them out of normalising falls in hospital, the Lead Nurse with a colleague developed a tool for staff to use to improve their care of patients at high risk of falling. The tool combined two patient safety techniques; a ‘checklist’ approach (Frank, 2006) and ‘intentional rounding’ (Owensboro Medical Health System, 2008). The checklist was a simple list of questions and actions (Figure 1) for staff to undertake with any patient who had been assessed as having a high risk of fallings (score >30). The ‘intentional rounding’ (IR) is the requirement to go round the ward and speak to those high risk patients every hour using the checklist. A simple checklist form was devised

for staff to use which then became part of the medical record for that patient. Using the PDSA methodology IR was tested by one nurse on one patient on one ward on one shift. Amendments to the checklist form were made before the tool was then disseminated to other nurses and then the whole ward team. A training pack was developed which helped the tool to be implemented in other ward areas within a short period of time.

In wards where a number of patients were assessed to be high risk then they were cohorted in a six bedded bay and a nurse allocated to that area all the time. This made the process of IR much quicker and focused staff resource towards the higher risk patients. The use of bed rails was also questioned and staff were reminded of the need to carefully risk assess and document their use. The checking of the bed rail status was then incorporated into the intentional rounding checklist. For many patients bed rails can increase the risk of harm from falling so should be used selectively.

There was some opposition to the introduction of IR. The concern was whether staff had the time to dedicate to IR given the many other priorities on a busy ward. The principle of PDSA was again employed and nurses realised that in some ways it helped them in prioritising and managing their workload. PDSA was also used to introduce IR to Care Assistants who provide much of the practical hands on care for patients. Registered nurses could then delegate IR to them and focus their attention on those tasks requiring their skills, such as dispensing medications.

The Lead Nurse set up weekly then monthly meeting to review progress with wards that implemented IR. As more wards took on the tool they were invited to the meeting. The reliability of the tool being used was tested informally by senior nurses visiting the wards and looking at the medical records of high risk patients.

Initial results from the first two wards to use IR (Wards A&B) were encouraging. The staff attitude towards patient falls changed, the number of falls reduced (Table 1) and there was a determination to learn from those incidents that did occur. The hospital incident form was supplemented by asking for further information as nurses suspected from studying the results, that patients with risk scores of between 20-30 and who had some form of cognitive impairment, could also benefit from IR.

The IR tool was spread rapidly from wards A&B to C-E and then within two months to a further four wards. Using the PDSA methodology what became evident during this rapid spread was that the training of staff was diluted and results were not sustained. Even on the initial wards (A&B) it was found that a patient had been assessed on seven occasions with a score greater than 30 but had not been put on IR because the nurses believed she was bed bound and therefore unlikely to fall. However, that patient did fall although was not injured. There was also a case of one patient who despite being on IR, fell a number of times and caused a 'blip' in the ward A results.

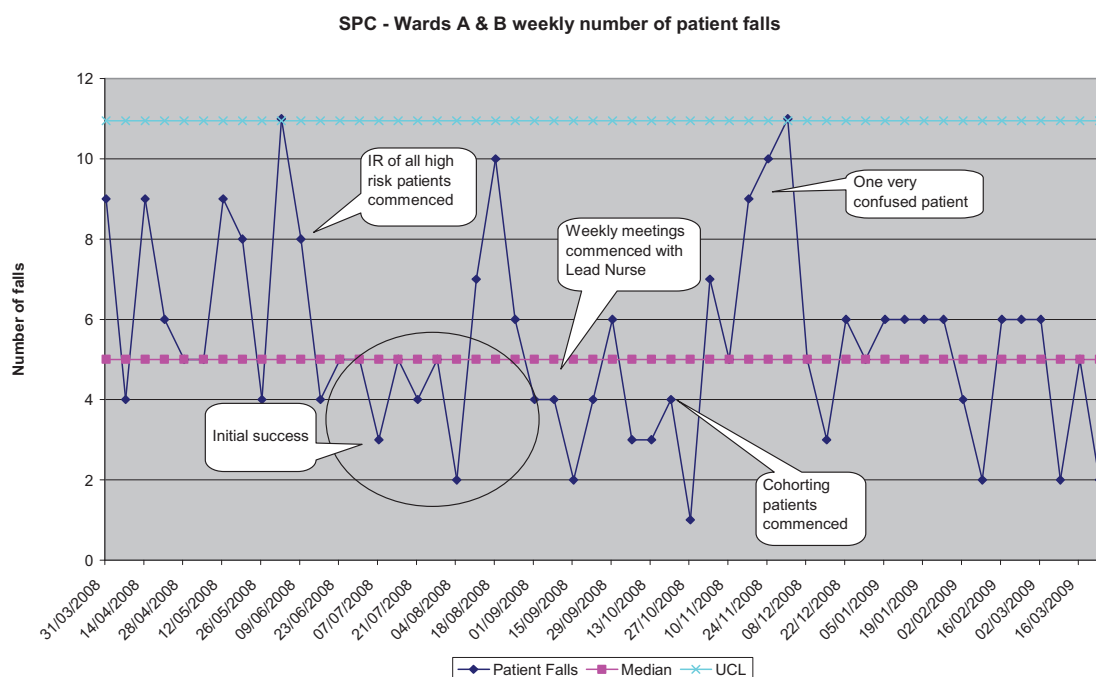


Table 1 Statistical process control chart - number of patient falls on wards A & B

Discussion

The ‘rebellion’ signifies the rejection of the ‘normal’ cultural goals; the rejection of the institutional means of acceptance and mechanical reporting of falls combined with minimal preventative action. Staff who engaged in the ‘rebellion’ put in place new goals (any fall is unacceptable) and institutional practice (staff using a checklist approach every hour to speak to patients at high risk of falling). It is accepted that to achieve no falls for patients in hospital is not a feasible goal. However, the unquestioning acceptance (normalisation) of the likelihood of falls in hospital results in a failure to implement systematic methods to reduce the risk for certain patients. The adoption of the changes required considerable leadership, education, training and persistent communication. However, even with the level of leadership given, the initial improvements were not sustained to the level desired. Using the PDSA methodology lessons were drawn as to why the initial improvement was not sustained as the new practice was spread. A number of changes to the training and auditing of results have since been put in place.

The findings support the theory of practical drift from a designed way of working (Snook, 2000) as set out in the ‘Falls Management Policy’. Staff drifted from the ‘Policy’ due to the competing priorities and finding practical ways to achieve their workload. In doing so they traded effectiveness for thoroughness (Hollnagel, 2004) and normalised their position as acceptable (Vaughan, 1996; Weick and Sutcliffe 2003). The danger of using a root case analysis approach to investigate those patient falls that result in significant harm, is that such a method often fails to look at the broader systemic issues such as practical drift and trade offs which are endemic amongst staff due to the competing pressures (Dekker, 2006). Senior managers can be left with the impression that the problem is isolated to particular wards rather than being a system wide issue that relates to more than one safety policy area.

Whilst PDSA is a powerful tool to achieve change, the cultural norms that have been established incrementally over many years, it takes a sustained effort by leaders to install rebellion. Regular meetings with staff were necessary plus constant feedback and learning when falls did occur. Those staff in this study who had been directly involved or felt responsible for a patient falling with a serious subsequent injury engaged with the change process at a much deeper level supporting the theory that ‘emotions can act as drivers or motivators, of subjects’ engagement with discourses’ (Garrety et al, 2003).

Conclusions

Improving patient safety in the midst of many competing demands requires considerable effort. Despite national initiatives and a revised Falls Management Policy staff in the case study hospital which was externally assessed as ‘excellent’, largely accepted the cultural norm that patient falls were part of being a patient in hospital. Theory relating to change in clinical practice to improve safety must take account of the strong cultural bias for staff to accept what is regarded as ‘normal’ and not challenge the status quo. Ways to encourage ‘rebellion’ include engaging with staff about the emotional stories of individual patients who suffer harm through a lack of proactive action.

Whilst there are limitations to this single case study there are aspects that have relevance to the wider practitioner interest in improving patient safety through cultural and practice-based change. From a theoretical perspective it seeks to build on the literature by developing the understanding of the drivers that reversed the normalisation of deviance and how safety was improved for patients in hospital.

References

- Bergman, M. M. Ed (2008), *Advances in Mixed Methods Research*, Sage, London.
- Creswell, J. W. (2003), *Research Design*, (2nd ed) Sage, London.
- Department of Health (2001), *The National Service Framework for Older People*, The Stationary Office, London.
- Dekker, S. (2006), *The Field Guide to Understanding Human Error*, Ashgate, Aldershot.
- De Vries, E.N. et al (2008), The incidence and nature of in hospital adverse events: a systematic review *Qual Saf Health Care*, 17:216-223
- Frank, C. (2006), Evidence based checklists for objective structured clinical examinations. *BMJ*, 333:417.
- Garrety K., Badman R., Morrigan V., Rifkin W., and Zanko M. (2003), The use of personality typing in organisational change : Discourse, emotions and reflexive subject. *Human Relations*, 56 ; 211-235.
- Grenier-Sennelier, C., I. Lombard, et al. (2002), Designing adverse event prevention programs using quality management methods: the case of falls in hospital. *Int J Qual Health Care*, 14(5): 419-426.
- Healey, F., Scobie, S., et al. (2008), Falls in English and Welsh hospitals: a national observational study based on retrospective analysis of 12 months of patient safety incident reports. *Qual Saf Health Care*, 17(6): 424-430.
- Healthcare Commission (2009), *Safely does It, implementing safer care for patients*, Commission for Healthcare Audit and Inspection, London.
- Hollnagel, E. (2004), *Barriers and Accident Prevention*, Ashgate, Aldershot.
- Langley, J., Nolan, K., Nolan, T., Norman, C. and Provost, L. (2006), *The Improvement Guide*, Jossey Bass, San Francisco, Ca.
- Merton, R.K. (1968), *Social Theory and Social Structure*, Free Press, New York.

- National Audit Office, (2005), *A Safer Place for Patients: Learning to improve patient safety*, Stationary Office, London.
- National Institute for Clinical Excellence (NICE) (2004), *Falls: the assessment and prevention of falls in older people*. Available at: <http://www.nice.org.uk/nicemedia/pdf/CG021NICEguideline.pdf> (14.3.09)
- Nightingale, P., Brady, T., Davies, A. and Hall, J. (2003), Capacity utilization revisited: software, control and the growth of large technical systems. *Industrial and Corporate Change*, Vol. 12 No. 3: 477-517.
- Oliver, D., Britton, M. et al. (1997), Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. *BMJ*, 315 (7115): 1049-1053.
- Oliver, D., Connelly, J.B., et al. (2007), Strategies to prevent falls and fractures in hospitals and care homes and effect of cognitive impairment: systematic review and meta-analyses. *BMJ*, 334 (7584): 82-88.
- Olsen, S., Neale, G., Schwab, K. et al (2007), Hospital staff should use more than one method to detect adverse events and potential adverse events. *Qual Saf Health Care*, 16:40-44.
- Owensboro Medical Health System, (2008), *Nursing Notes*. Available at: <http://omhs.org/docs/newsletter/September%202008.pdf> (4.2.09)
- Sari, AB-A., Sheldon, T.A. and Cracknell, A. (2007), Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care*, 16:434-9.
- Snook, S.A. (2000), *Friendly Fire*, Princeton University Press, Princeton.
- Stuart, I., McCutcheon, D., Handfield, R., Mclachlin, R. and Samson, D. (2002), Effective case research in operations management - a process perspective. *Journal of Operations Management*, Vol. 20 No. 5, pp. 419-33.
- Vaughan, D. (1996), *The Challenger Launch Decision*, Chicago University Press, Chicago.
- Waring, J.J. (2005) Beyond blame: cultural barriers to medical incident reporting. *Social Science & Medicine*, 60, 1927-1935.
- Waring, J., Harrison, S., and McDonald, R. (2007), A culture of safety or coping? Ritualistic behaviours in the operating theatre. *Journal of Health Services Research and Policy*, 12 (Suppl 1):3-9.
- Weick, K. and Sutcliffe, K. (2003), Hospitals as Cultures of Entrapment: A Re-Analysis of the Bristol Royal Infirmary. *California Management Review*, Vol. 48 No. 2:73-84.
- Wheeler, D. J. (2003), *Making Sense of Data*, SPC Press, Knoxville.
- Yin, R.K. (2003), *Case Study Research, Design and Methods*, Sage, London.

Publication 4

Sweeney, K. & Williams, M. D. (2010) Safety, systems, complexity, and resilience: What makes organizations safe? In Haxby, E., Hunter, D. & Jaggar, S. (Eds.) *An Introduction to Clinical Governance and Patient Safety*. Oxford, Oxford University Press.

Chapter 51

Safety, systems, complexity, and resilience: What makes organizations safe?

Kieran Sweeney and Michael D. Williams

Key points

- ◆ Healthcare organizations are complex adaptive systems consisting of inter-related sub-systems of micro-communities of individuals.
- ◆ Emergent properties of healthcare systems can be explained by complexity theory.
- ◆ Individuals make errors but good teams can reduce their number and impact.
- ◆ Resilience refers to an organization's ability to maintain its systems whilst adapting to external pressures.

Introduction

In this chapter it is proposed that when viewing the activities of a hospital, it is helpful to see that organization as a complex adaptive system (CAS). Complex organizations are constructed of inter-relating subunits (like departments), and those subunits, in turn, consist of micro-communities of individuals (like doctors and nurses). The reciprocal interaction of these micro-communities co-creates the culture of these subunits and, by scaling up, the culture of the organization itself. In this chapter, healthcare organizations are viewed as complex adaptive systems, requiring the application of systems thinking, and safety is seen as an emergent property of the inter-relationship of subunits and the individuals within them.

Safety, governance, and emergent properties

Safety is part of clinical governance, which was rather opaquely defined by Scally and Donaldson in 1998 as:

a framework through which NHS organisations are accountable for continuously improving the quality of their services, and safeguarding high standards of care, by creating an environment in which excellence in clinical care will flourish.

Seven pillars of clinical governance were identified – clinical effectiveness, risk management, patient experience, communication effectiveness, resource effectiveness,

strategic, and learning effectiveness. Once established however, these domains tended to spawn silos of enthusiastic people scurrying about doing their best to spread the word. NHS organizations became governance aware, acutely so in some cases, when visited by the Health Commission, but not much excellence was flourishing. It became apparent that it was the inter-relationship between these domains that could lead to a culture more conducive to excellence, but to achieve this, leadership, teamwork, good, clear communication, and above all systems awareness were needed.

It became apparent that the pillars should not be seen as separate domains at all – this was like thinking about the weather, and considering temperature, pressure, humidity, and local geography separately, whereas each of these parameters, although measurable independently, is inter-dependent on all the others.

For healthcare, it is crucial to see safety as an emergent property of the interaction of all the seven domains of clinical governance, co-creating a quality environment. To grasp this idea theoretically, an understanding of systems, and their complexity and adaptability is required.

Healthcare organizations as complex adaptive systems

Early management approaches to strategy in the NHS regarded organizations as machines operated from the centre and, more importantly, predictable in their responses and outputs. The introduction of systems thinking into NHS management in the 1990s changed this. In systems theory, one accepts that a system (here, an organization like a hospital) consists of a collection of subunits, which interact with one other to produce an output that is greater (and more complex) than the outputs of the individual subunits. Theoretically, there are four features of systems that help us consider how they operate. These are shown in Table 51.1.

While these four features are not mutually exclusive, it is easy to see that hospitals are better thought of as soft, pluralistic organizations generating a range of problems, some of which can be solved algorithmically, and others which need a different kind of more organic management approach.

Over the last decade, complexity theory has been increasingly used to describe how large healthcare organizations work. Complexity theory helps our understanding of the challenges inherent in managing such organisations, and why they frequently fail. The key ideas underpinning complexity in organizational theory are organic: organizations grow, they need to be fertilised with creative conversations, their roots lie in

Table 51.1 Characteristics of systems

Feature of the system	Comment
Hard	Expressed in a rigorous, usually numerical manner
Soft	More ambiguous, fluid and evolving, less easily defined numerically
Positivist	Permitting one undisputed interpretation of events
Pluralistic	Accepts that there are multiple views, each legitimately held within the system

the continuous interaction of all the people who work in them, and the pattern of their behaviour marks the 'culture' of the organization – in what is termed its emergent properties.

Complexity is one of four ways in which a system evolves dynamically (and here a 'system' means a set of agents, their interaction, and sense of purpose). A 'system' can be understood to evolve, dynamically, in four ways.

Systems can be:

1. Static – there is no dynamic evolution
2. Ordered – exhibiting rhythmic predictable dynamic movement, like a combustion engine
3. Chaotic – appearing to change randomly, but predicated on hidden rhythms and patterns, which can be computed mathematically, like the weather.
4. Complex – complex systems operate in the state of dynamic evolution between order and chaos, like the tube of the surfer's wave, before it crashes (chaotically) onto the beach.

A complex adaptive system (CAS) has five stages, or activities:

1. Receptive context
2. Complex responsive processes
3. Self-organization
4. Adaption or co-evolution
5. Emergent properties

For a system to evolve creatively, the agents in it needs to have first a *receptive context* – a general set of values the agents understand and accept. Then, the agents need to interact continuously in both predictable and unpredictable ways, in what is termed a series of *complex responsive processes*. The patterning of these processes gives rise to self-organization, essentially the patterns of behaviour by which the organization can recognize itself. The most often quoted example of self-organizing behaviour is when birds form a flock – flocking is self-organizing behaviour. Clearly, organizations don't just self-organize within their own boundaries, they will interact with other organizations in their environment, and change some of their activities accordingly. In complexity theory, this is called *adaption or co-evolution*, the fourth principle of complexity. In turn, the patterning of these self-organizing and adaptive activities gives rise to the 'feel' or 'culture' of the organization, called the organization's *emergent properties*; if flocking is the self-organizing behaviour, then the flock is the emergent property. By the same token, a wave is an emergent property of water. Emergence is a higher order feature of complex adaptive systems, and has been defined as

'the arising of novel and coherent structures, patterns and properties during the process of self-organization in complex systems'.

It therefore refers to the potential within such systems to create properties that could not have been predicted by understanding the nature of each separate element in the system.

It is proposed that, in thinking about safety, errors, near misses, or beneficial patient outcomes in a hospital, it is helpful to see that organization as a complex adaptive system (CAS). A blame culture often fails to see an NHS organisation as complex; individuals are not isolated free-standing agents who just slot into a system somewhere. They are embedded in one system, for example their ward nursing group, which in turn forms part of a larger group, the intensive care department, which again exists within a particular directorate of the hospital, and so on. These units in the organization are interconnected, the agents engage in a complex responsive process continuously, and in so doing self-organize themselves (into rotas, for example), and interact with others (e.g. managers in the system). The output of all the interaction, within and between all the levels of the system, will be the emergent property of the system. This could be called the 'feel' or 'culture' of the hospital – what it feels like to work there.

Safety, safety culture, and high-reliability organizations (HRO)

Safety culture is hard to define. Safety expert Charles Vincent suggests we summon up our own experience in organizations, like hospitals, and remember wards in which the atmosphere was relaxed, but where the staff were conscientious, and standards uniformly high. Equally, we can recall more slapdash wards, where risks were run, and potentially dangerous practices might have been developing beneath the surface. Vincent asks us to consider the 'culture' of such wards, and suggests

culture points to the powerful influence of social forces in moulding our behaviour; we are all more malleable than we like to think and to some extent we develop good or bad habits according to the prevailing ethos around us.

From the explanation of CAS above, it should be clear that the 'moulding of our behaviour' will take place via the complex responsive processes, and that the 'ethos' of the organization will be one of its emergent properties.

Ever since the 1940s high-reliability organizations (HROs), like aviation or the nuclear industry, have developed a systems approach to error. HROs accept that errors are inevitable, and are likely to be the outcome of a systematic failure within the wider organization, rather than the blameworthy behaviour of a single person. Many of the lessons learned in the aviation industry, where most accidents are attributable to human factors, have been usefully transferred to clinical practice. Although Vincent himself warns against seeing the parallels too pervasively, there is no doubt that looking at how HROs learn from errors has been helpful to some key areas of medical practice. This is especially important where clinical professionals work in close-knit teams, as in operating theatres. Teams, like individuals can create or erode safety. A landmark study by Lingard in 2004 concluded that about 25% of all communications in an operating theatre could be classified as communication failures, either unclear, incomplete, or just plain wrong. Working poorly, a team multiplies the possibility of error; working well, the team creates an environment which is safer than the combined efforts of the individuals; safety becomes, in those circumstances, a defining property of the system.

Error detection and management have always been central features of training programmes for cockpit crews in civil aviation. Human vulnerability to stressors, the nature of error, and how to respond to it are studied to achieve three aims:

1. Good teamwork
2. A culture of problem solving through good communication.
3. Reducing the possibility that the crew will make errors

Key to this is that the aviation industry has a culture, which accepts

- ◆ that errors are inevitable;
- ◆ that systems failures or weaknesses often contribute to them, and
- ◆ that being open about the proneness to error, by fostering good communication, is an essential way of reducing their incidence.

Unfortunately this contrasts with the culture within medical practice. This difference between the cultures of the two industries was highlighted in a study by Helmreich in 2000. Junior cockpit crew were asked if they should be able to question decisions made by senior colleagues; unanimously they said yes, and that this was, quite simply, another defence against error. When posed the same question, one quarter of senior consultant surgeons said that junior surgeons should not question their seniors.

The issue of a hierarchy of communication is not simply of theoretical or academic importance. In 2001, Wayne Jowett died after he had been given an intrathecal injection of vincristine, a drug that should only be given intravenously. The subsequent investigation showed a myriad of tacit assumptions about the roles of the two junior doctors involved, which, while not completely unreasonable, led to an interaction between the two, which was literally fatal. In particular, when the more junior doctor was asked why he did not challenge the registrar about the second, fatal syringe, which contained vincristine, but was referred to as methotrexate by the registrar, he said, 'I was a junior doctor, and did what I was told by the registrar ... I assumed he had the knowledge ... and I did not intend to challenge him.' When sentenced to 8 months imprisonment, the registrar said, 'I know it's a lame excuse, but I am a human being'. Everybody is, which is precisely why solid systems, good teamwork, and open communication are needed to protect us from our own fallibility.

Resilience

The idea of resilience is a key element in understanding why NHS organizations have a poor patient safety record, and understanding how that situation might improve.

Resilience is a proactive concept of safety. In healthcare, it refers to an organization's ability to absorb and adapt to increasing pressure, to prevent adverse patient outcomes in the face of near disruption. In resilience, an operating envelope is imagined, within which the system is designed to be competent. Speculation then occurs as to how the system can maintain such competence in the face of pressures, such as rising demand for emergency care, which move it away from this safe operating envelope. Four boundaries to the operating envelope have been proposed:

1. Financial failure
2. Target failure

3. Unacceptable working conditions
4. Failure of safety.

There are obvious managerial and political pressures on healthcare organisations to achieve financial balance, and at the same time reduce waiting times. This promotes a strong productivity culture, which has to be balanced against proper working arrangements for key staff, like junior doctors. These three influences act interdependently, continuously challenging the fourth, safety boundary.

To promote and maintain an acceptable safety culture, and to manage risk appropriately, healthcare organizations have to know where their operating point lies in relation to the boundaries of this safe operating envelope. Recent work demonstrates that NHS organizations are predominantly influenced by the first two of the four boundaries described; financial and target failure. When organizations are seen as organic, flexible and adaptive, it is clear that in order to 'protect' these two boundaries, the other two (working practices and safety) may come under increasing pressure. Work is currently underway to develop more robust metrics to describe all four boundaries to help balance the dominant influence of the measures used to monitor the finance and targets boundaries.

Summary

Humans are fallible, and error is inevitable. Safe organisations accept that errors will occur, and try to reduce their number and effect by building systems to protect the people working within them.

Complexity theory helps us to understand how organizations operate, how they are simultaneously predictable and unpredictable, and how the interaction of the people working within them is fundamental to constructing the ethos, or culture of the organization.

The notion of resilience allows us to imagine how organisations evolve within a safe operating envelope, and what pressures impact on such organisations to move it out of this safe buffer zone. Clinicians need clear communication, good teamwork, and a supportive infrastructure to work safely.

Further reading

- Vincent C (2006) *Patient Safety*. Churchill Livingstone: London.
- Flin R, O'Connor P, Crichton M. (2008) *Safety at the Sharp End: A Guide to Non-Technical Skills*. Ashgate: Aldershot.
- Sweeney K (2006) *Complexity in Primary Care: Understanding its Value*. Radcliffe Medical Press: Abingdon.
- Williams M (2008) A Descriptive Model of Resilience in NHS Hospitals. Proceedings of Euroma Conference, 16–18 May, Groningen: Holland.
- Kernick D (ed) (2004) *Complexity and Healthcare Organisations*. Radcliffe Medical Press: Abingdon.